EDITOR'S NOTE

Welcome to the first Issue of Volume 75 of *The Federal Communications Law Journal*. We are the nation's premiere communications law journal and the official journal of the Federal Communications Bar Association. This year, as we celebrate the 75th anniversary of this publication, we look forward to sharing articles and student Notes that showcase the range of issues relevant to the field of technology and communications law.

To start, this first Issue provides thoughtful scholarship on topics including influencer marketing, data privacy, artificial intelligence, and the evolution of First Amendment jurisprudence.

This Issue begins with an examination of how best to protect journalists if the Supreme Court acts on recent calls to reconsider the landmark defamation case *New York Times Co. v. Sullivan*. In this article, Matthew L. Schafer, Adjunct Professor of Law at Fordham University School of Law, and Jeff Kosseff, Associate Professor of Cybersecurity Law at the United States Naval Academy, argue for a federal Freedom of Speech and Press Act that would codify the protections established in *Sullivan*.

This Issue also features four student Notes. In the first Note, Delaney Dunn dives into the ever-growing world of influencer marketing. She argues that influencers should be held accountable for fraudulently inflating their followings and deceiving their brand partners.

The second Note, written by Thompson Hangen, focuses on the privacy risks involved with a Central Bank Digital Currency. He argues that, while establishing a Central Bank Digital Currency would strengthen the U.S. dollar, it necessitates an expansion of federal financial data privacy laws.

The third Note, written by Dallin Albright, explores the potential consequences of automated journalism. Albright identifies gaps in libel law regarding news content generated with artificial intelligence. He argues the existing negligence standard is best suited to addressing instances of libel resulting from algorithmic speech.

Finally, our last student Note in the first Issue was authored by Robin Briendel. She proposes a four-step test for determining when school administrators possess the authority to regulate student speech occurring off-campus.

The Editorial Board of Volume 75 would like to thank the FCBA and The George Washington University Law School for their continued support of the *Journal*. We also appreciate the hard work of the authors and editors who contributed to this Issue.

The Federal Communications Law Journal is committed to providing its readers with in-depth coverage of relevant communication law topics. We welcome your feedback and encourage the submission of articles for publication consideration. Please direct any questions or comments about this Issue to fclj@law.gwu.edu. Articles can be sent to fcljarticles@law.gwu.edu. This Issue and our archive are available at http://www.fclj.org.

Julia Dacy
Editor-in-Chief

Federal Communications Law Journal

The *Federal Communications Law Journal* is published jointly by the Federal Communications Bar Association and The George Washington University Law School. The *Journal* publishes three issues per year and features articles, student Notes, essays, and book reviews on issues in telecommunications, the First Amendment, broadcasting, telephony, computers, Internet, intellectual property, mass media, privacy, communications and information policymaking, and other related fields.

As the official journal of the Federal Communications Bar Association, the *Journal* is distributed to over 2,000 subscribers, including Association members as well as legal practitioners, industry experts, government officials and academics. The *Journal* is also distributed by Westlaw, Lexis, William S. Hein, and Bloomberg Law and is available on the Internet at www.fclj.org.

The *Journal* is managed by a student Editorial Board, in cooperation with the Editorial Advisory Board of the FCBA and two Faculty Advisors.

Federal Communications Bar Association

The Federal Communications Bar Association (FCBA) is a volunteer organization of attorneys, engineers, consultants, economists, government officials and law students involved in the study, development, interpretation, and practice of communications and information technology law and policy. From broadband deployment to broadcast content, from emerging wireless technologies to emergency communications, from spectrum allocations to satellite broadcasting, the FCBA has something to offer nearly everyone involved in the communications industry. That's why the FCBA, more than two thousand members strong, has been the leading organization for communications lawyers and other professionals since 1936.

Through its many professional, social, and educational activities, the FCBA offers its members unique opportunities to interact with their peers and decision-makers in the communications and information technology field, and to keep abreast of significant developments relating to legal, engineering, and policy issues. Through its work with other specialized associations, the FCBA also affords its members opportunities to associate with a broad and diverse cross-section of other professionals in related fields. Although the majority of FCBA members practice in the metropolitan Washington, D.C. area, the FCBA has eleven active regional chapters, including: Atlanta, Carolina, Florida, Midwest, New England, New York, Northern California, Southern California, Pacific Northwest, Rocky Mountain, and Texas. The FCBA has members from across the U.S., its territories, and several other countries.

The George Washington University Law School

Established in 1865, The George Washington University Law School (GW Law) is the oldest law school in Washington, D.C. The Law School is accredited by the American Bar Association and is a charter member of the Association of American Law Schools. GW Law has one of the largest curricula of any law school in the nation with more than 275 elective courses covering every aspect of legal study.

GW Law's home institution, The George Washington University, is a private institution founded in 1821 by charter of Congress. The Law School is located on the University's campus in the downtown neighborhood familiarly known as Foggy Bottom.

The *Federal Communications Law Journal* is published by The George Washington University Law School and the Federal Communications Bar Association three times per year. Offices are located at 2028 G Street NW, Suite LL-020, Washington, D.C. 20052. The *Journal* can be reached at fclj@law.gwu.edu, and any submissions for publication consideration may be directed to fcljarticles@law.gwu.edu. Address all correspondence with the FCBA to the Federal Communications Bar Association, 1020 19th Street NW, Suite 325, Washington, D.C. 20036-6101.

Subscriptions: Subscriptions are $30 per year (domestic), $40 per year (Canada and Mexico), and $50 per year (international). Subscriptions are to be paid in U.S. dollars, and are only accepted on a per-volume basis, starting with the first issue. All subscriptions will be automatically renewed unless the subscriber provides timely notice of cancellation. Address changes must be made at least one month before publication date, and please provide the old address or an old mailing label. Please direct all requests for address changes or other subscription-related questions to the journal via email at fclj@law.gwu.edu.

Single and Back Issues: Each issue of the current volume can be purchased for $15 (domestic, Canada and Mexico) or $20 (international), paid in U.S. dollars. Please send all requests for single or back issues to fclj@law.gwu.edu.

Manuscripts: The *Journal* invites the submission of unsolicited articles, comments, essays, and book reviews mailed to the office or emailed to fcljarticles@law.gwu.edu. Manuscripts cannot be returned unless a self-addressed, postage-paid envelope is submitted with the manuscript.

Production: The citations in the *Journal* conform to THE BLUEBOOK: A UNIFORM SYSTEM OF CITATION (Columbia L. Rev. Ass'n et al. eds., 21st ed., 2021). Variations exist for purposes of clarity and at the editors' discretion. The *Journal* is printed by Joe Christensen, Inc.

Citation: Please cite this issue as 75 FED. COMM. L.J. ___ (2022).

FEDERAL COMMUNICATIONS LAW JOURNAL

THE TECH JOURNAL

GW LAW VOLUME 75 ISSUE 1 fcba THE TECH BAR

NOVEMBER 2022

ARTICLES

Until 1964, states were free to penalize journalists, activists, and others for criticizing the most powerful figures in the United States. That changed with the Supreme Court's opinion in *New York Times Co. v. Sullivan*, which requires public officials suing for defamation to establish actual malice, a daunting hurdle. Over the next three decades, the Court expanded on *Sullivan* and built a framework that provides vital First Amendment protections for modern journalism, online commentary, and other criticism. Those safeguards face their greatest threats ever, as high-profile figures weaponize defamation lawsuits and two Supreme Court justices call on their colleagues to join them in reconsidering *Sullivan*. As the Supreme Court has recently demonstrated, it will not shy away from rethinking even the most vital and established constitutional protections. To prevent the damage to free speech caused by a sudden reversal of *Sullivan*, we propose the federal Freedom of Speech and Press Act, which codifies many of the protections of *Sullivan* and its progeny and preempts state defamation laws that do not satisfy certain minimum standards that preserve "uninhibited, robust, and wide-open" debate across the country.

NOTES

Influencer marketing is the future of the advertising industry, and it does not appear to be a promising one. Marketers sought out influencers to avoid the shortfalls of traditional marketing but instead found an entirely new set of concerns. Influencers are able to exploit the present system for their own personal gain with little regard for the companies they are hurting and little concern for the repercussions of their actions. They are paid based on likes and follows on their social media pages, and they regularly falsely inflate these numbers to steal money out of the pockets of the brands they are dealing with.

standard used for libel claims by private individuals. This would assign to operators and creators of autonomous journalism software a reasonable duty of care to ensure that the information they publish is accurate and non-defamatory by following industry procedures for promoting journalistic accuracy. The unique nature of artificial intelligence as it applies to speech requires that courts adapt their existing legal standards to match the challenges presented by new technologies.

Here, There, and Everywhere: Defining the Boundaries of the "Schoolhouse Gate" in the Era of Virtual Learning

The Internet and social media have caused a dramatic change in how today's students communicate. With virtual learning remaining prominent as the COVID-19 pandemic endures, and students continuing their near-constant use of phones and computers, the line of what constitutes "on-campus" activity is blurry at best. This lack of clarity has significantly complicated the ability of school officials and courts to determine what speech is outside the scope of schools' disciplinary authority.

This Note evaluates Supreme Court precedent concerning the regulation of student speech, the circuit courts' differing approaches to tackling off-campus student speech, as well as the Supreme Court's recent decision, *Mahanoy Area School District v. B.L.*, addressing the scope of schools' authority to punish students for speech generated off-campus on social media. Ultimately, this Note concludes that the ill-defined standards concerning the ability of school officials to discipline students for off-campus speech create massive amounts of uncertainty and problems for students, schools, and the courts. This Note suggests that the Supreme Court should articulate a uniform mode of assessment for school administrators and courts to use for determining whether a student's speech is entitled to First Amendment protection.

Protecting Free Speech in a Post-*Sullivan* World

Matthew L. Schafer & Jeff Kosseff*

TABLE OF CONTENTS

* Matthew L. Schafer is an Adjunct Professor of Law, Fordham University School of Law. Jeff Kosseff is an Associate Professor of Cybersecurity Law, United States Naval Academy. The views expressed in this piece are only the authors' and do not represent Fordham University School of Law, the Defense Department, Department of Navy, the Naval Academy, or any other party.

I. INTRODUCTION

Thomas Cooper, who Thomas Jefferson classed as "the greatest man in America," once said that "[t]he doctrine of libel is, in all countries, a doctrine of power."[1] So it remains today. Today, the wealthy, famous, and otherwise powerful regularly resort to libel threats and libel lawsuits not to redress a cognizable injury to their reputation but instead to silence and punish their critics and make to-be critics think twice before speaking. Luckily, the U.S. Supreme Court has recognized in three decades of case law that the First Amendment displaces much of the common law of libel (and other speech-based torts), making it harder for tech billionaires, Hollywood elites, and political partisans to weaponize libel law.

Starting in 1964, at the height of the civil rights movement, the Supreme Court in *New York Times Co. v. Sullivan* said for the first time that libel lawsuits brought by public officials must be considered against the backdrop of our "profound national commitment to the principle that debate on public issues should be uninhibited, robust, and wide-open," despite that such debate "may well include vehement, caustic, and sometimes unpleasantly sharp attacks."[2] *Sullivan* transformed the common law by placing the burden under the First and Fourteenth Amendments on public officials to prove falsity and a heightened fault standard called "actual malice."[3] That standard requires a public official plaintiff to plead and prove that the defendant published the allegedly defamatory statement knowing that it was false or with a high degree of awareness of its probable falsity.[4] The Court's recognition in *Sullivan* was hailed as an occasion for "dancing in the streets."[5] It was "a great case" when it was decided and is, today, a landmark precedent.[6]

Sullivan and the cases that came after it, however, hang in the balance now more than ever before. We have not seen libel plaintiffs flock to courts in such numbers since the 1980s, "a time of growing libel litigation, of enormous judgments and enormous costs."[7] And, even despite *Sullivan*, several plaintiffs still manage to succeed. Short of a jury verdict in their favor, libel plaintiffs can measure their success in years-long defense costs that can easily exceed $1-2 million depending on the case. For plaintiffs seeking retribution more than redress, putting a defendant through the time and trouble is well worth the squeeze.

While this might suggest that *Sullivan* should be shored up, or perhaps that the Supreme Court should recognize other protections under the First

---

1. James C. Humes, *The Nation's First Civil Disobedient*, 58 AM. BAR ASS'N J. 259, 259 (1972).
2. N.Y. Times Co. v. Sullivan, 376 U.S. 254, 270 (1964).
3. *Id.* at 279-80.
4. *Id.*; *see also* Garrison v. Louisiana, 379 U.S. 64, 74 (1964).
5. Harry Kalven, Jr., *The New York Times Case: A Note on "The Central Meaning of the First Amendment,"* 1964 SUP. CT. REV. 191, 221 n.125 (1964) (quoting Alexander Meiklejohn).
6. Anthony Lewis, New York Times v. Sullivan *Reconsidered: Time to Return to "the Central Meaning of the First Amendment,"* 83 COLUM. L. REV. 603, 603 (1983).
7. *Id.*

Amendment, some on the Court have called for overruling *Sullivan*. Clarence Thomas was first: "The constitutional libel rules adopted by this Court in [*Sullivan*] and its progeny broke sharply from the common law of libel, and there are sound reasons to question whether the First and Fourteenth Amendments displaced this body of common law."[8] He has twice renewed this call.[9] Nor is he alone. Neil Gorsuch, in 2021, joined him, suggesting that *Sullivan* might be the problem, not the solution.[10] And although she has not made her position known recently, as a law professor in the 1990s Elena Kagan pondered whether the Court had "extended the *Sullivan* principle too far."[11]

Sullivan may not be reversed next term or five terms on. But having seen the scramble to protect bodily autonomy in the wake of the Court overturning *Roe v. Wade*, the time to protect landmarks like *Sullivan* is now.[12] Here, we argue that Congress should take up and pass a preemption statute. This proposed statute would set baseline national standards, some previously adopted by the Court as a constitutional matter and others only ever considered by it, that must be satisfied to maintain a defamation action based on interstate speech. By doing so, Congress could insulate the press and the public from fallout that will follow in the wake of overruling *Sullivan*. This approach has the added benefit of not establishing a national law of libel nor a new procedural scheme such as an anti-SLAPP, both of which are more ambitious proposals that we think have low likelihood of gaining traction in Congress no matter how appropriate such approaches might be.

On our way to proposing this statutory scheme, we first review *Sullivan* itself and the sociopolitical environment in which the Court decided that case before we turn to some of the cases that followed it. This review is necessary to understand the import of the statutory language we aim to propose. We next examine recent calls to revisit *Sullivan*. To explain why such rethinking is dangerous, we provide an overview of the increasing weaponization of the law of libel by all sorts of plaintiffs, proving that there is a real, emergent problem that Congress can address by adopting our proposal. We then discuss statutory preemption of the state law of libel, using Section 230 of the Communications Decency Act as a model. Finally, we propose statutory language to protect freedom of speech and of the press and discuss how we arrived at this language.

---

8. McKee v. Cosby, 139 S. Ct. 675, 678 (2019) (Thomas, J., concurring in denial of certiorari).

9. Berisha v. Lawson, 141 S. Ct. 2424, 2424-25 (2021) (Thomas, J., dissenting from denial of certiorari); Coral Ridge Ministries Media, Inc. v. S. Poverty L. Ctr., 142 S. Ct. 2453, 2454-55 (2022) (Thomas, J., dissenting from denial of certiorari).

10. *Berisha*, 141 S. Ct. at 2429 (Gorsuch, J., dissenting from denial of certiorari).

11. Elena Kagan, *A Libel Story:* Sullivan *Then and Now*, 18 LAW & SOC. INQUIRY 197, 205 (1993) (reviewing ANTHONY LEWIS, MAKE NO LAW: THE SULLIVAN CASE AND THE FIRST AMENDMENT (1991)).

12. *See, e.g.*, Dobbs v. Jackson Women's Health Org., 142 S. Ct. 2228, 2305 (2022).

II. *NEW YORK TIMES CO. V. SULLIVAN*

It was about a month after the Greensboro Four refused to leave the "Whites Only" lunch counter.[13] On March 29, 1960, the *Times* ran an advertisement titled *Heed Their Rising Voices*.[14] The ad, paid for by Committee to Defend Martin Luther King and the Struggle for Freedom in the South, was intended to throw a spotlight on young civil rights protesters "engaged in widespread non-violent demonstrations in positive affirmation of the right to live in human dignity as guaranteed by the U.S. Constitution and the Bill of Rights."[15] These demonstrations were met with "an unprecedented wave of terror" detailed in the ad "by those who would deny and negate that document."[16]

That ad did not name a single police officer in Alabama, and the *Times* distributed just 394 daily copies of the newspaper in that state—a paltry amount relative to its circulation of 650,000 copies.[17] Nevertheless, L.B. Sullivan, a member of the Commissioners of the City of Montgomery and in that role supervisor of the police, sued the *Times* over the ad, arguing that its references to "police" could be read to refer to him specifically.[18] There was also a companion case, *Abernathy v. Sullivan*, that has receded from memory but proves that *Sullivan* was not merely a case about freedom of the press.[19] Rather, it implicated freedom of speech for the individual too, as Sullivan also sued four black ministers, Ralph David Abernathy, S.S. Seay Sr., Fred L. Shuttlesworth, and J.E. Lowery, whose names appeared on the advertisement without their permission.[20]

The ad was not without its issues. While it reported that protesters sang *My Country, 'Tis of Thee* on the state capitol steps, in fact they sang the national anthem.[21] While it reported that the dining hall had been padlocked, in fact the university denied entry to certain students because they did not have dining tickets.[22] Moreover, while it reported that the police ringed the campus, in fact they deployed near the campus.[23] While nine students had been expelled, it was not because they led a demonstration at the Capitol, but because they demanded to be served at a lunch counter.[24] And while the ad

13. Michael Ray et al., *Greensboro Sit-in*, ENCYC. BRITANNICA, https://www.britannica.com/event/Greensboro-sit-in [https://perma.cc/B4SF-XNDY] (last visited Nov. 6, 2022).
14. N.Y. Times Co. v. Sullivan, 376 U.S. 254, 257 (1964).
15. *Id.* at 256.
16. *Id.*
17. *Id.* at 260 n.3.
18. *Id.* at 258.
19. *See generally id.*
20. William E. Lee, *Citizen-Critics, Citizen Journalists, and the Perils of Defining the Press*, 48 GA. L. REV. 757, 758-59 (2014).
21. *Sullivan*, 376 U.S. at 258-59.
22. *Id.* at 259.
23. *Id.*
24. *Id.*

stated that Martin Luther King, Jr. had been arrested seven times, in fact he had only been arrested four times.[25]

At trial, Sullivan put on evidence that he had not been involved in the misconduct as alleged in the ad.[26] Instead, he argued that much of the conduct pre-dated his time as commissioner of the police.[27] He made no effort to prove actual damages and instead relied on witness testimony from a former employer that had they believed the ad, they would have been less likely to associate with him.[28] The judge instructed the jury that the statements were libelous *per se* and not privileged.[29] He also told the jury that because the statements were *per se* libelous, Sullivan did not have to put on evidence of actual damage.[30] Falsity and malice, he told the jury, were also presumed.[31] Finally, he told the jury that punitive damages need not have any relation to actual damages.[32] The jury then found for Sullivan, awarding him $500,000.[33] The Alabama Supreme Court affirmed.[34]

On January 7, 1963, the Supreme Court granted certiorari, citing "the importance of the constitutional issues involved" as to both the *Times* and the individual defendants.[35] In a unanimous opinion, authored by Justice William Brennan, it reversed.[36]

At first, the Court summarized the outlines of Alabama's libel law. A statement was libelous *per se* where "the words 'tend to injure a person . . . in his reputation' or to 'bring [him] into public contempt.'"[37] When it came to a public official, a finding that the statement "'injure[d] him in his public office, or impute[d] misconduct to him in his office, or want of official integrity, or want of fidelity to a public trust'" satisfied that standard.[38] Where the plaintiff was a public official, "his place in the governmental hierarchy" was "sufficient evidence" that "statements that reflect" on government reflect on those in charge of it.[39] Thereafter, the defendant was left with no defense unless he could show that the charge is "true in all [its] particulars."[40] Moreover, absent a showing of truth, "general damages are presumed, and may be awarded without proof of pecuniary injury."[41] To get punitive

25. *Id.*
26. *Id.*
27. *Sullivan*, 376 U.S. at 259.
28. *Id.* at 260.
29. *Id.* at 262.
30. *Id.*
31. *Id.*
32. *Id.*
33. *Sullivan*, 376 U.S. at 262.
34. *Id.* at 256.
35. *Id.* at 264; *see also* N.Y. Times Co. v. Sullivan., 371 U.S. 946, 946 (1963).
36. *Sullivan*, 376 U.S. at 264.
37. *Id.* at 267.
38. *Id.*
39. *Id.*
40. *Id.*
41. *Id.*

damages, however, the plaintiff "apparently" had to show malice.[42] Neither "good motives" nor "belief in truth" negated a finding of malice.[43]

Turning to whether the Constitution had anything to say about this state of affairs, the Court said it was required to consider Sullivan's case "against the background of a profound national commitment to the principle that debate on public issues should be uninhibited, robust, and wide-open, and that it may well include vehement, caustic, and sometimes unpleasantly sharp attacks on government and public officials."[44] The Sedition Act of 1789, the Court wrote, "first crystallized a national awareness of the *central meaning* of the First Amendment."[45] That statute prohibited publishing "any false, scandalous and malicious writing or writings against the government of the United States, or either house of the Congress . . . , or the President . . . , with intent to defame."[46] Violators risked a fine of $5,000 and up to five years in jail.[47] Unlike at common law, the statute permitted defendants a defense of truth and, nominally, placed in the hands of the jury both law and fact.[48]

According to the Court, the statute had been forced through the Federalist-controlled Congress keen on keeping John Adams in power, the Court noted that it "was vigorously condemned as unconstitutional in an attack joined in by Jefferson and Madison" in the Virginia and Kentucky resolutions.[49] As adopted by the Virginia General Assembly, the Virginia resolution said that the Sedition Act authorized the national government to exercise "'a power not delegated by the Constitution, but, on the contrary, expressly and positively forbidden by one of the amendments.'"[50] The power authorized by the Act "ought to produce universal alarm, because it is levelled against the right of freely examining public characters and measures, and of free communication among the people."[51]

The Court then observed that Madison, who drafted the First Amendment, had viewed the Sedition Act as unconstitutional and harmful to a republican government.[52] In that government, Madison had said, "The people, not the government, possess the absolute sovereignty."[53] The colonists distrusted "power itself at all levels," but especially "concentrated power."[54] Importantly, the government established by the Founders was

42. *Sullivan*, 376 U.S. at 264.
43. *Id.*
44. *Id.* at 270 (citing Terminiello v. City of Chicago, 337 U.S. 1, 4 (1949); De Jonge v. Oregon, 299 U.S. 353, 365 (1937)).
45. *Id.* at 273 (emphasis added).
46. *Id.* at 273-74.
47. *Id.* at 273.
48. *Sullivan*, 376 U.S. at 274.
49. *Id.*
50. *Id.* at 274 (quoting James Madison, *Madison's Report on the Virginia Resolutions*, *in* 4 ELLIOT'S DEBATES: THE DEBATES IN THE SEVERAL STATE CONVENTIONS ON THE ADOPTION OF THE FEDERAL CONSTITUTION 554 (Jonathan Elliot ed., 1876) [hereinafter *Madison's Report*]).
51. *Id.* (quoting *Madison's Report*, *supra* 50, at 554).
52. *Id.* at 274-76; *see also* Matthew L. Schafer, *In Defense: New York Times v. Sullivan*, 82 LA. L. REV. 81, 91, 137 (2021).
53. *Sullivan*, 376 U.S. at 274 (quoting *Madison's Report*, *supra* 50, at 569).
54. *Sullivan*, 376 U.S. at 274.

"'altogether different' from its British form, under which the Crown was sovereign."[55] It was, thus, "necessary" in America to have "a different degree of freedom . . . of the press."[56] As Madison had said on the floor of Congress years earlier, "If we advert to the nature of Republican Government, we shall find that the censorial power is in the people over the Government, and not in the Government over the people."[57]

Historically, the Court concluded that the People in fact exercised that power. Madison had written, "In every state, probably, in the Union, the press has exerted a freedom in canvassing the merits and measures of public men, of every description, which has not been confined to the strict limits of the common law."[58] In this country, he added, "On this footing the freedom of the press has stood; on this foundation it yet stands."[59] Thus, it was "manifestly impossible," consistent with the Constitution, "to punish the intent to bring those who administer the government into disrepute or contempt, without striking at the right of freely discussing public characters and measures."[60] From this, the Court found that "[t]he right of free public discussion of the stewardship of public officials was, thus, in Madison's view, a fundamental principle of the American form of government."[61]

Although the Sedition Act expired on its own after Jefferson took office so its constitutionality had never been considered by the Court, the Court wrote that "the attack upon its validity has carried the day in the court of history."[62] Jefferson pardoned those convicted, finding that the Act was "a nullity, as absolute and as palpable as if Congress had ordered us to fall down and worship a golden image."[63] The fines issued under it were repaid and, as time wore on, one politician observed that "its invalidity was a matter 'which no one now doubts.'"[64] It was no surprise that other justices had also drawn into question the validity of the act, including Oliver Wendell Holmes, Louis Brandeis, Robert Jackson, and William O. Douglas.[65]

The ad in Sullivan's case, targeted as it was at the government, "would seem clearly to qualify for the constitutional protection" in light of this history.[66] The only question, the Court explained, was whether that protection

---

55. *Id.* at 274 (quoting *Madison's Report*, *supra* 50, at 569).
56. *Id.* at 275 (quoting *Madison's Report*, *supra* 50, at 570).
57. *Id.* (quoting 4 ANNALS OF CONG. 934 (1855)).
58. *Id.* (quoting *Madison's Report*, *supra* 50, at 570).
59. *Id.* (quoting *Madison's Report*, *supra* 50, at 570).
60. *Sullivan*, 376 U.S. at 275 n.15 (quoting *Madison's Report*, *supra* 50, at 575) ("The value and efficacy of this right depends on the knowledge of the comparative merits and demerits of the candidates for public trust, and on the equal freedom, consequently, of examining and discussing these merits and demerits of the candidates respectively.").
61. *Id.*
62. *Id.* at 276.
63. *Id.* (quoting Letter from Thomas Jefferson to Abigail Adams (July 22, 1804), *in* 4 THE WORKS OF THOMAS JEFFERSON 555, 555-56 (H.A. Washington ed., 1884)).
64. *Id.* (quoting S. REP. NO. 122, at 3 (1836)).
65. *Id.* (citing Abrams v. United States, 250 U.S. 616, 630 (1919) (Holmes, J., dissenting); Beauharnais v. Illinois, 343 U.S. 250, 288-89 (1952); WILLIAM O. DOUGLAS, THE RIGHT OF THE PEOPLE 47 (1958)).
66. *Sullivan*, 376 U.S. at 271.

was "forfeit[ed]" because of the "falsity of some of its factual statements and by its alleged defamation of respondent."[67] It found that it was not.[68]

As to falsity, the Court said that it had "consistently refused to recognize an exception for any test of truth . . . and especially one that puts the burden of proving truth on the speaker."[69] As Madison also had written, "Some degree of abuse is inseparable from the proper use of every thing; and in no instance is this more true than in that of the press."[70] Consistent with this, the Court observed that "erroneous statement is inevitable in free debate, and that it must be protected if the freedoms of expression are to have the 'breathing space' that they 'need . . . to survive.'"[71] In the end, cases meant to "impose liability for erroneous reports of the political conduct of officials reflect the obsolete doctrine that the governed must not criticize their governors."[72]

The Court then found that the First and Fourteenth Amendments required "a federal rule that prohibits a public official from recovering damages for a defamatory falsehood relating to his official conduct unless he proves that the statement was made with 'actual malice.'"[73] Actual malice, it wrote, equated to "knowledge that [the statement] was false or with reckless disregard of whether it was false or not."[74] In other words, public officials would have to prove that the defendant published a calculated falsehood in order to recover damages.

The Kansas Supreme Court adopted a "like rule" in 1908 in *Coleman v. MacLennan*.[75] There, a politician sued a newspaper that charged him with mismanagement.[76] In adopting that rule, the Kansas court noted that it "is of the utmost consequence that the people should discuss the character and qualifications of candidates for their suffrages."[77] The importance of those kinds of discussions to democracy "more than counterbalance the inconvenience of private persons whose conduct may be involved."[78] In such a system, "occasional injury to the reputations of individuals must yield to the public welfare."[79]

There was also a symmetry to the rule, as it was "analogous to the protection accorded a public official when he is sued for libel by a private citizen."[80] All States at that time accorded privileges to statements made by

67. *Id.*
68. *Id.* at 271-73.
69. *Id.* (citing Speiser v. Randall, 357 U.S. 513, 525-26 (1958)).
70. *Id.* (quoting James Madison, quoting *Madison's Report*, *supra* 50, at 571).
71. *Id.* at 271-72 (quoting NAACP v. Button, 371 U.S. 415, 433 (1963)).
72. *Sullivan*, 376 U.S. at 272 (quoting Sweeney v. Patterson, 128 F.2d 457, 458 (D.C. Cir. 1942)).
73. *Id.* at 279-80.
74. *Id.* at 280.
75. *Id.* (citing Coleman v. MacLennan, 98 P. 281, 281-82 (Kan. 1908)).
76. *Id.*
77. *Id.* at 281 (quoting *Coleman*, 98 P. at 286).
78. *Sullivan*, 376 U.S. at 281 (quoting *Coleman*, 98 P. at 286).
79. *Id.* (quoting *Coleman*, 98 P. at 286).
80. *Id.* at 282.

public officials in their duties "unless actual malice can be proved."[81] Otherwise, "the threat of damage suits would . . . 'inhibit the fearless, vigorous, and effective administration of policies of government' and 'dampen the ardor of all but the most resolute, or the most irresponsible, in the unflinching discharge of their duties.'"[82] Mirroring that, the Court found that a similar privilege should apply to "the citizen-critic of government," because it was "*as much his duty to criticize as it is the official's duty to administer*."[83] In a republican government, "It would give public servants an unjustified preference over the public they serve, if critics of official conduct did not have a fair equivalent of the immunity granted to the officials themselves."[84]

Having adopted the actual malice rule, the Court then applied it, anticipating that Sullivan would seek a new trial that would be as unfair as the last.[85] As to the individual defendants, the question was easy, as Sullivan introduced "no evidence whatever that they were aware of any erroneous statements or were in any way reckless in that regard."[86] As to the *Times*, the question required more thought, but ultimately, the answer was the same. First, testimony demonstrated that the *Times* believed the ad to be true.[87] Second, the failure to retract was not evidence of actual malice because the *Times* did not even believe the ad was about Sullivan.[88] Third, the allegation that clips in the *Times*' archives refuted facts in the ad thereby demonstrating a calculated falsehood was also insufficient because those at the *Times* responsible for the ad were unaware of those clips.[89]

The verdict was "constitutionally defective" in another way: "it was incapable of supporting the jury's finding that the allegedly libelous statements were made 'of and concerning'" Sullivan.[90] First, the ad never mentioned Sullivan by name or position.[91] Several statements alleged to be defamatory did not even relate to the police, let alone Sullivan.[92] As to the statements that police ringed the campus or that Dr. King had been arrested seven times, the Court found that "[a]lthough the statements may be taken as referring to the police, they did not on their face make even an oblique reference to respondent as an individual."[93] None of the witness testimony stated any reason to believe Sullivan was involved beyond the mere association with the police.[94] Were that alone sufficient to render the statements actionable, it would violate the rule that "prosecutions for libel on

81. *Id.*
82. *Id.* (quoting Barr v. Matteo, 360 U.S. 564, 571 (1959)).
83. *Id.* (emphasis added).
84. *Sullivan*, 376 U.S. at 282-83.
85. *See id.* at 284-85.
86. *Id.* at 286.
87. *Id.*
88. *Id.*
89. *Id.* at 287.
90. *Sullivan*, 376 U.S. at 288.
91. *Id.*
92. *Id.*
93. *Id.* at 289.
94. *Id.*

government" have no "place in the American system of jurisprudence."[95] Permitting recovery "would sidestep this obstacle by transmuting criticism of government, however impersonal it may seem on its face, into personal criticism, and hence potential libel, of the officials of whom the government is composed."[96]

Although *Sullivan* was unanimous, three Justices believed that the Court should provide even stronger protections to defamation defendants. Justice Hugo Black, joined by Justice William Douglas, wrote that the First Amendment provides the press with "an absolute immunity for criticism of the way public officials do their public duty."[97] Likewise, Justice Arthur Goldberg, also joined by Douglas, wrote that "the First and Fourteenth Amendments to the Constitution afford to the citizen and to the press an absolute, unconditional privilege to criticize official conduct despite the harm which may flow from excesses and abuses."[98] These three believed that even public officials who could establish actual malice should be unable to sue for defamation.

In the end, *Sullivan* embraced the argument of philosopher Alexander Meiklejohn that the First Amendment is necessary to foster self-governance.[99] Indeed, a year after he wrote *Sullivan*, Justice Brennan delivered a lecture at Brown University in which he explicitly linked *Sullivan* to Meiklejohn's philosophy: "The first amendment question was whether its protections nevertheless limit a state's power to apply traditional libel law principles, since the statements were made in criticism of the official conduct of a public servant."[100] "In other words, the case presented a classic example of an activity that Dr. Meiklejohn called an activity of 'governing importance' within the powers reserved to the people and made invulnerable to sanctions imposed by their agency-governments."[101]

Of course, *Sullivan*, despite being a unanimous opinion, was never preordained. Thirty years after it was decided, Anthony Lewis, who wrote a biography of *Sullivan* in his book *Make No Law*, posed the contrary result: "Suppose that Southern judges and juries had had the last word, that the press had no higher recourse in the American system."[102] The proposition requires little imagination. By using libel law as a political weapon, the Southern judicial system could have controlled the narrative and suppressed the rising civil rights movement.

Before trial, the *Times* even struggled to find an Alabama lawyer to represent it in the face of outrage "whipped up" against the *Times* by the political establishment in Alabama.[103] When the *Times* New York lawyer

95. *Id.* at 291 (quoting City of Chicago v. Tribune Co., 139 N.E. 86, 88 (Ill. 1923)).
96. *Sullivan*, 376 U.S. at 292.
97. *Id.* at 295 (Black, J., concurring).
98. *Id.* at 298 (Goldberg, J., concurring).
99. William J. Brennan, Jr., *The Supreme Court and the Meiklejohn Interpretation of the First Amendment*, 79 HARV. L. REV. 1, 14 (1965).
100. *Id.*
101. *Id.*
102. ANTHONY LEWIS, MAKE NO LAW: THE SULLIVAN CASE AND THE FIRST AMENDMENT 36 (1991).
103. *Id.* at 24.

traveled to Alabama in preparation for the case, he stayed at a hotel under an assumed name.[104] Once the case got to trial, it was assigned to Judge Walter Jones, a "devotee of the Confederacy and the Southern way of life."[105] Jones would later say that the case would be tried not under the Fourteenth Amendment but according to "white man's justice."[106] He empaneled an all-white jury.[107] And while the transcript of the trial referred to white lawyers with the honorific "Mr.", for the Black lawyers, the transcript read only "Lawyer Crawford" or "Lawyer Seay" as they were, according to racist custom, undeserving of the "Mr."[108]

For the political establishment in Alabama, Sullivan's lawsuit, and those that followed, were wildly successful.[109] As Lewis recounted, the day after the jury verdict in *Sullivan*, the *Alabama Journal* published an editorial arguing that the verdict would "have the effect of causing reckless publishers of the North . . . to make a re-survey of their habit of permitting anything detrimental to the South and its people to appear in their columns."[110] *Sullivan* had "changed the rules": "The Times was summoned more than a thousand miles to Montgomery to answer for its offense. . . . The only way to prevent such long distance summons is to print the truth."[111]

As Lewis observed though, after *Sullivan*, printing the truth was far from an easy thing to do: "The rules applied by Judge Jones made it forbiddingly difficult to write anything about the realities of Southern racism in the 1960's without risking heavy damages for libel."[112] That was, of course, the whole point. Sullivan, and other public officials, "were out to transform the traditional libel action, designed to repair the reputation of a private party, into a state political weapon to intimidate the press."[113] The purpose was "to discourage not false but true accounts of life under a system of white supremacy," making it impossible to write about lynching, segregation, and the rest of the South's cruel history.[114]

At the time, the $500,000 verdict against the *Times* was the largest ever libel judgment in Alabama,[115] and more would come in the tag-along suits, totaling $3 million.[116] There was a question if the *Times* could survive litigation over the ad, to say nothing of the other lawsuits then pending across the South brought by public officials against Northern agitators.[117] As Lewis explained, "By the time the Supreme Court decided the *Sullivan* case, in 1964, Southern officials had brought nearly $300 million in libel actions against the

104. *Id.*
105. *Id.* at 25.
106. *Id.* at 26.
107. *Id.* at 27.
108. LEWIS, *supra* note 101, at 27.
109. *Id.* at 34.
110. *Id.*
111. *Id.*
112. *Id.* at 34.
113. *Id.* at 35.
114. LEWIS, *supra* note 101, at 35.
115. *Id.*
116. *Id.*
117. *Id.*

press."[118] Libels lawsuits had become the weapon of choice to "repress[] the movement for civil rights."[119]

III. THE PROGENY

The same year the Court decided *Sullivan*, Harry Kalven, Jr. wrote: "It is not easy to predict what the Court will see in the [*Sullivan*] opinion as the years roll by."[120] But, he added, "the invitation to follow a dialectic progression from public official to government policy to public policy to matters in the public domain, like art, seems to me to be overwhelming."[121] He was right; from 1964 to the early 1990s, the Court continued to tinker with the balance between the sanctity with which the law treated one's reputation and the constitutional rights of freedom of speech and of the press, sometimes suggesting that it would tilt that balance in favor of reputation and sometimes tilting it in favor of speech. *Sullivan*'s progeny is well documented extensively elsewhere and is only repeated in brief here.[122]

Garrison v. Louisiana. Just months after the Court decided *Sullivan*, it considered the constitutionality of Louisiana's criminal libel law. In *Garrison v. Louisiana*, Jim Garrison, the district attorney of Orleans Parish, made several disparaging statements about criminal court judges in the Parish.[123] In substance, he accused those judges of "inefficiency, laziness, and excessive vacations."[124] As a result, the State charged him with criminal defamation, and a judge convicted him.[125]

The Court first considered whether its decision in *Sullivan*, a civil case, should be extended to the criminal context. In finding that it should, the Court explained that there was "no merit in the argument that criminal libel statutes serve interests distinct from those secured by civil libel laws, and therefore should not be subject to the same limitations."[126] In fact, the Court wrote, by the first half of the nineteenth century, civil libel actions had already begun to replace the use of criminal libel laws.[127] In other words, they served the same purpose—to suppress unpopular speech.

It then considered whether the common law defense of truth and good motives could be incorporated into the First Amendment as a constitutional protection in criminal cases. The question was relevant as the Louisiana statute at issue allowed a conviction based on a true statement where that statement was made with ill-will.[128] The Court found that the common law

118. *Id.* at 36.
119. *Id.* at 35.
120. Kalven, *supra* note 5, at 221.
121. *Id.*
122. *See* LEE LEVINE & STEPHEN WERMIEL, THE PROGENY: JUSTICE WILLIAM J. BRENNAN'S FIGHT TO PRESERVE THE LEGACY OF NEW YORK TIMES V. SULLIVAN (2014), for the comprehensive history of *Sullivan*'s progeny.
123. Garrison v. Louisiana, 379 U.S. 64, 65 (1964).
124. *Id.* at 66.
125. *Id.* at 65.
126. *Id.* at 67.
127. *Id.* at 68-69.
128. *Id.* at 71-72.

defense was insufficient, holding that the requirement that a defendant show truth *and* good motives was too burdensome.[129] Instead, it explained, "where the criticism is public officials and their conduct of public business, the interest in private reputation is overborne by the larger public interest, secured by the Constitution, in the dissemination of truth" irrespective of motives.[130]

Finally, even as to false statements, the Court found that *Sullivan* prevented the imposition of criminal liability so long as those statements were not calculated falsehoods. As the Court put it, "even where the utterance is false, the great principles of the Constitution which secure freedom of expression in this area preclude attaching adverse consequences to any except the knowing or reckless falsehood" consistent with *Sullivan*.[131] Indeed, "[f]or speech concerning public affairs is more than self-expression; it is the essence of self-government."[132]

Curtis Publishing Co. v. Butts; Associated Press v. Walker. Both *Sullivan* and *Garrison* had the analog to the Sedition Act as both cases reflected liability for speech—either civil or criminal—for criticizing public officials in performance of their public functions. *Curtis Publishing Co. v. Butts* and the companion case, *Associated Press v. Walker*, would mark the first major expansion of *Sullivan*—and they would do so without a majority opinion.[133]

These cases forced the Court to consider the foretold conflict recognized by Kalven as to the application of *Sullivan* to "persons who are not public officials, but who are 'public figures' and involved in issues in which the public has a justified and important interest."[134] *Sullivan*, Justice John Marshall Harlan noted, had "expressly reserved" what the "sweep" of its logic may be as to this question.[135] The question now had to be answered, however, because of a "sharp division" among lower courts as to the import of *Sullivan* outside the context of public official plaintiffs.[136]

Wally Butts was the athletic director of the University of Georgia, and in that role, had been accused of trying to fix football games.[137] While Georgia was a state school, the Georgia Athletic Association, a private entity, employed Butts.[138] Butts was "well-known" at the time and had been the football coach for Georgia.[139] He sued the newspaper that had accused him of fixing the games, and before the Court had decided *Sullivan*, a jury awarded him nearly half a million dollars.[140] The Fifth Circuit affirmed, although one

129. *Garrison*, 379 U.S. at 72-73.
130. *Id.*
131. *Id.* at 73.
132. *Id.* at 74-75.
133. Curtis Publ'g Co. v. Butts, 388 U.S. 130 (1967) (plurality opinion).
134. *Id.* at 134.
135. *Id.* (citing N.Y. Times Co. v. Sullivan, 376 U.S. 254, 283 n.23 (1964)).
136. *Id.* (citing Clark v. Pearson, 248 F. Supp. 188, 194 (D.D.C. 1965) (stating that *Sullivan* only applied to "officials in the high echelons")). *See id.* at 134 n.1 for the Court's list of lower court decisions that contributed to the division.
137. *Id.*
138. *Id.*
139. *Butts*, 388 U.S. at 134 (plurality opinion).
140. *Id.* at 138.

judge, consistent with *Sullivan* and *Garrison*, would have reversed as the instruction may have "allow[ed] recovery on a showing of intent to inflict harm or even the culpably negligent infliction of harm, rather than the intent to inflict harm through falsehood."[141]

Edwin Walker was a racist who had "a long . . . career in the United States Army before resigning to engage in political activity."[142] When the *Associated Press* published a dispatch accusing him of encouraging violent opposition to the desegregation at the University of Mississippi, Walker was no longer in the Army but maintained a political following as a private person.[143] Walker sued in Texas (of all places), and the jury awarded him $800,000.[144] The trial judge vacated the punitive damages award, which reduced the verdict by $300,000 on the grounds that Walker failed to establish actual malice.[145] He refused to vacate the balance, asserting that "[t]ruth alone" was a sufficient defense and there was no compelling public policy reason to extend *Sullivan*.[146] On appeal, the Texas Court of Civil Appeals affirmed, and the Supreme Court of Texas denied further review.[147]

While Harlan announced the judgment of the Court, it was Chief Justice Earl Warren's opinion for himself that controlled. Parting with Harlan and three other Justices, Warren found that public figures, like public officials, must also plead and prove that a libel defendant acted with actual malice. Separately joined by Justices Black, Douglas, Brennan, and Byron White, Warren said that while he agreed with the result in Harlan's opinion, he disagreed with its failure to extend *Sullivan* to public figures.[148] Warren looked at the case through a pragmatic lens: public figures' "views and actions with respect to public issues and events are often of as much concern to the citizen as the attitudes and behavior of 'public officials' with respect to the same issues and events."[149]

Differentiating between public officials and public figures in American society had "no basis in law, logic, or First Amendment policy," Warren wrote.[150] Lines between "governmental and private sectors [were] blurred" in 1960s America.[151] Policy determinations that had historically been wholly government were now "channeled through formal political institutions are now originated and implemented through a complex array of boards, committees, commissions, corporations, and associations, some only loosely connected with the Government."[152] Since the 1930s, there had been "a rapid fusion of economic and political power, a merging of science, industry, and government, and a high degree of interaction between the intellectual,

141. *Id.* at 139-40.
142. *Id.* at 140.
143. *Id.*
144. *Id.* at 141.
145. *Butts*, 388 U.S. at 141-42 (plurality opinion).
146. *Id.* at 142.
147. *Id.*
148. *Id.* at 162 (Warren, C.J., concurring).
149. *Id.*
150. *Id.* at 163.
151. *Butts*, 388 U.S. at 163 (Warren, C.J., concurring).
152. *Id.*

governmental, and business worlds."[153] All the while, power had become "much more organized" in the "private sector."[154]

A similar blurring between public officials and public figures attended this transformation. Many "who do not hold public office at the moment are nevertheless intimately involved in the resolution of important public questions or, by reason of their fame, shape events in areas of concern to society at large."[155] While they were not born of the political process, they were *a part of* that process.[156] As a result, the citizenry had "a legitimate and substantial interest in the conduct of such persons, and freedom of the press to engage in uninhibited debate about their involvement in public issues and events is as crucial as it is in the case of 'public officials.'"[157] In a way, that public figures "are not amenable to the restraints of the political process only underscores the legitimate and substantial nature of the interest, since it means that public opinion may be the only instrument by which society can attempt to influence their conduct."[158] Thus, Warren found that "public men," generally, must prove actual malice.

St. Amant v. Thompson. In *St. Amant v. Thompson*, the Supreme Court again reviewed an opinion by the Louisiana Supreme Court. This time, it was a civil case related to a statement made during a speech by a candidate for public office.[159] The issue, though, was narrow: whether the state court had appropriately applied the test for actual malice.[160] While the state high court had found sufficient evidence that the defendant made the statement with "reckless disregard" as to its truth, the Court reversed, finding that the state court had treated the inquiry as an objective one rather than subjective one.[161] As the Court explained, "reckless conduct is not measured by whether a reasonably prudent man would have published, or would have investigated before publishing."[162] Instead, the Court said that there "must be sufficient evidence to permit the conclusion that the defendant in fact entertained serious doubts as to the truth of his publication."[163]

Rosenbloom v. Metromedia, Inc. *Rosenbloom* was a defamation lawsuit based on news reporting of the arrest of a nudist magazine purveyor for distributing obscene materials.[164] All agreed that "the police campaign to enforce the obscenity laws was an issue of public interest" and that the magazine purveyor was neither a public official nor public figure.[165] The only question was whether, as a private individual, the plaintiff nevertheless had

153. *Id.*
154. *Id.*
155. *Id.*
156. *Id.*
157. *Butts*, 388 U.S. at 164 (Warren, C.J., concurring).
158. *Id.*
159. St. Amant v. Thompson, 390 U.S. 727, 728 (1968).
160. *Id.*
161. *Id.* at 730.
162. *Id.* at 731.
163. *Id.*
164. Rosenbloom v. Metromedia, Inc., 403 U.S. 29, 33-34 (1971) (plurality opinion), *abrogated by* Gertz v. Robert Welch, Inc., 418 U.S. 323 (1974).
165. *Id.* at 40.

to plead and prove actual malice as the statement was about his "involvement in an event of public or general interest."[166] Affirming the Third Circuit, which found that actual malice must be shown, Brennan announced the judgment of the Court, but he lacked a majority.

According to Brennan, "Self-governance in the United States presupposes far more than knowledge and debate about the strictly official activities of various levels of government."[167] Instead, "[o]ur efforts to live and work together in a free society not completely dominated by governmental regulation necessarily encompass far more than politics in a narrow sense."[168] As a result, he argued that the First Amendment "if it would fulfill its historic function in this nation, must embrace all issues about which information is needed or appropriate to enable the members of society to cope with the exigencies of their period."[169]

Sullivan and those cases that followed, however, had focused only on the status of plaintiff and not the underlying controversy. This created an "artificiality" in the public's interest in any given case between "'public' and 'private' individuals or institutions."[170] A matter of public interest though did not become less so simply because a private figure was involved.[171] On the contrary, the "public's primary interest is in the event; the public focus is on the conduct of the participant and the content, effect, and significance of the conduct, not the participant's prior anonymity or notoriety."[172] The case before the Court demonstrated as much: whether the plaintiff was a private figure was irrelevant to the public's weightier interest in ensuring that criminal conduct was pursued appropriately.[173] This was, Brennan argued, the import of the Court's prior decisions, even though they spoke in terms of a plaintiff's status as a public or private individual.[174]

While Brennan's opinion was joined by Chief Justice Warren and Justice Harry Blackmun, others concurred only in judgment. Black concurred, consistent with his long-held belief that "the First Amendment does not permit the recovery of libel judgments against the news media even when statements are broadcast with knowledge they are false."[175] White also concurred only in judgment. For White, his colleagues were trying to do much in a case that required only a little. *Sullivan*, he wrote, "made clear that discussion of the official actions of public servants such as the police is constitutionally privileged."[176] Because official conduct is often targeted at private figures, *Sullivan* necessarily allowed for the intrusion upon the privacy or reputations of "private citizens against whom official action is directed."[177]

166. *Id.* at 31-32.
167. *Id.* at 41.
168. *Id.*
169. *Id.*
170. *Rosenbloom*, 403 U.S. at 41 (plurality opinion).
171. *Id.* at 43.
172. *Id.*
173. *Id.*
174. *Id.*
175. *Id.* at 57 (Black, J., concurring).
176. *Rosenbloom*, 403 U.S. at 61 (White, J., concurring).
177. *Id.*

It gave "the press the right not only to censure and criticize officials but also to praise them and the concomitant right to censure and criticize their adversaries," like the magazine purveyor.[178] Thus, he would have recognized "a privilege to report and comment upon the official actions of public servants in full detail, with no requirement that the reputation or the privacy of an individual involved in or affected by the official action be spared from public view."[179]

Justices John Marshall Harlan II, Thurgood Marshall, and Potter Stewart dissented. While they recognized that the case implicated the First and Fourteenth Amendments, they thought that Brennan's opinion would constitutionalize too much of the state law of libel. Instead, Harlan would have held "unconstitutional, in a private libel case, jury authority to award punitive damages," which he said was "unconfined by the requirement that these awards bear a reasonable and purposeful relationship to the actual harm done."[180] Marshall and Stewart would have taken a narrower view on permissible liability, arguing that damages should be limited to actual losses and otherwise leaving standards of liability to the states so long as strict liability was not imposed.[181]

Gertz v. Robert Welch, Inc. *Gertz*, like *Rosenbloom*, presented the question of "whether a newspaper or broadcaster that publishes defamatory falsehoods about an individual who is neither a public official nor a public figure may claim a constitutional privilege against liability for the injury inflicted by those statements."[182] Richard Nucio, a Chicago police officer, shot and killed Ronald Nelson.[183] Nelson's family retained Elmer Gertz to represent them in litigation against Nuccio.[184]

Around the same time, the far-right John Birch Society was publishing articles warning of a propaganda war against law enforcement.[185] As part of that effort, it published an article, "FRAME-UP: Richard Nuccio And The War On Police."[186] That article reported that testimony at Nuccio's criminal trial was false and part of the "Communist campaign against the police."[187] Although Gertz had little involvement in the criminal trial, the article fingered him as the mastermind of the "frame-up," reported that he had a criminal file so big it would take "'a big, Irish cop to lift,'" and said he was an official of the "Marxist League for Industrial Democracy."[188]

Gertz sued the John Birch Society.[189] On a motion for summary judgment, the defendant invoked *Sullivan*, arguing that Gertz was either a

178. *Id.*
179. *Id.* at 62.
180. *Id.* at 77 (Harlan, J., dissenting).
181. *Id.* at 86 (Marshall, J., dissenting).
182. Gertz v. Robert Welch, Inc., 418 U.S. 323, 332 (1974).
183. *Id.* at 325.
184. *Id.*
185. *Id.*
186. *Id.* at 325-26.
187. *Id.* at 326.
188. *Gertz*, 418 U.S. at 326.
189. *Id.* at 325, 327.

public official or a public figure, but the court concluded that he was not.[190] At trial, the jury awarded Gertz $50,000.[191] The court, however, had a change of heart post-verdict and found that *Sullivan* did apply and that Gertz had to establish actual malice.[192] It did so not because Gertz was a public figure but because *Sullivan* reached "discussion of any public issue without regard to the status of a person defamed therein."[193] The Seventh Circuit affirmed based on Brennan's intervening plurality decision in *Rosenbloom*.[194]

The Supreme Court reversed and rejected *Rosenbloom*. Justice Lewis Powell, writing for the Court, began by recognizing the Court's struggle to "define the proper accommodation between the law of defamation and the freedoms of speech and press protected by the First Amendment."[195] After a long review of *Sullivan* and its progeny, the Court began on "common ground."[196] While it questioned the constitutional value of false statements, it explained that such statements are "inevitable in free debate."[197] Punishing such errors risked "inducing a cautious and restrictive exercise of the constitutionally guaranteed freedoms of speech and press."[198] Thus, in *Sullivan* and elsewhere, the Court had held that the First Amendment requires "we protect some falsehood in order to protect speech that matters."[199]

On the other side of the ledger was the state interest in compensating citizens whose reputations had been unwarrantedly sullied. An individual's right to his or her reputation "'reflects no more than our basic concept of the essential dignity and worth of every human being—a concept at the root of any decent system of ordered liberty.'"[200] Thus, "'some antithesis between freedom of speech and press and libel actions persists, for libel remains premised on the content of speech and limits the freedom of the publisher to express certain sentiments, at least without guaranteeing legal proof of their substantial accuracy.'"[201]

Rather than pick a side between these two competing interests, the Court sought a middle ground. The media, it wrote, "[is] entitled to act on the assumption that public officials and public figures have voluntarily exposed themselves to increased risk of injury from defamatory falsehood concerning them," but '[n]o such assumption is justified with respect to a private individual."[202] Because a private figure has not "relinquished" her interest in her reputation to the public because of her conduct, she "has a more

190. *Id.*
191. *Id.* at 328-29.
192. *Id.* at 329.
193. *Id.*
194. *Gertz*, 418 U.S. at 329.
195. *Id.* at 325.
196. *Id.* at 339.
197. *Id.* at 340.
198. *Id.*
199. *Id.* at 341.
200. *Gertz*, 418 U.S. at 341 (quoting Rosenblatt v. Baer, 383 U.S. 75, 92 (1966) (Stewart, J., concurring)).
201. *Id.* at 342 (quoting Curtis Publ'g Co. v. Butts, 388 U.S. 130, 152 (1967) (plurality opinion)).
202. *Id.* at 345.

compelling call on the courts for redress of injury inflicted by defamatory falsehood."[203]

Still, First Amendment concerns required some limitations. The Court held that "so long as they do not impose liability without fault, the States may define for themselves the appropriate standard of liability."[204] Thus, it allowed private figures to recover under whatever fault standard is set by a state but also gave media defendants some breathing space by not allowing recovery under a strict liability standard. The Court's conclusion was not based on the "belief that the considerations which prompted the adoption of the" actual malice rule in *Sullivan* and *Curtis Publishing Co.* "are wholly inapplicable to the context of private individuals."[205] Rather, it was the strength and legitimacy of the States' countervailing interest in protecting private figures that required a more nuanced approach.[206]

That interest, however, did not extend to providing for presumed or punitive damages.[207] The need to limit presumed damages was necessary because libel is an "oddity of tort law" that allowed for recovery of damages "without any proof that such harm actually occurred."[208] The risk of rogue juries assessing catastrophic damages "unnecessarily compound[ed] the potential of any system of liability for defamatory falsehood to inhibit the vigorous exercise of First Amendment freedoms."[209] As importantly, allowing juries uncontrolled discretion made it likely that they would "punish unpopular opinion rather than to compensate individuals for injury sustained."[210] And, the States had "no substantial interest in securing for plaintiffs such as this petitioner gratuitous awards of money damages far in excess of any actual injury."[211] The Court then held that "defamation plaintiffs who do not prove knowledge of falsity or reckless disregard for the truth" could only recover actual damages.[212]

Although Blackmun had joined the plurality opinion in *Rosenbloom*, he concurred in *Gertz*.[213] Despite the "illogical" retreat from *Rosenbloom*, Blackmun joined in *Gertz* for two reasons.[214] First, he was satisfied that the limits on presumed and punitive damages "eliminate[d] significant and powerful motives for self-censorship that otherwise are present in the traditional libel action."[215] Second, he thought it vital to provide certainty in the law: "I feel that it is of profound importance for the Court to come to rest in the defamation area and to have a clearly defined majority position that

203. *Id.*
204. *Id.* at 347.
205. *Id.*
206. *Gertz*, 418 U.S. at 347.
207. *Id.* at 348-49.
208. *Id.* at 349.
209. *Id.*
210. *Id.*
211. *Id.*
212. *Gertz*, 418 U.S. at 349.
213. *Id.* at 353 (Blackmun, J., concurring).
214. *Id.*
215. *Id.* at 354.

eliminates the unsureness engendered by *Rosenbloom*'s diversity."[216] Had his vote not been needed, he would have followed *Rosenbloom*.[217]

Douglas wrote again to express his and Black's view that the Court should get out of the business of defining boundaries to First Amendment freedoms where the text of that Amendment allowed for none.[218] He noted that the First Amendment barred Congress from passing any civil libel law, as Thomas Jefferson had observed in 1798.[219] Nor had Congress ever done so.[220] While Congress had passed the Sedition Act, as the Court observed in *Sullivan*, the "general consensus was that the Act constituted a regrettable legislative exercise plainly in violation of the First Amendment."[221] His point was simple: if Congress lacked authority to pass either civil or criminal libel laws under the First Amendment, the States lacked any authority to do so under the Fourteenth.[222]

Maintaining his unbending position that the First Amendment did not allow any exceptions, Douglas said that the sanction of jury damages in civil libel cases "impinge[d] upon free and open discussion."[223] This was especially the case because speech that "arouses little emotion is little in need of protection," while speech that is "marked by highly charged emotions" may become "a virtual roll of the dice separating them from liability for often massive claims of damage."[224] Whether it be negligence or actual malice, Douglas feared that the Court's ever "proliferating standards in the area of libel" were likely to increase self-censorship.[225]

Brennan dissented.[226] True, he explained, the majority held that the First Amendment did act as a limit even on libel actions brought by private figures involved in a matter of public interest.[227] This reflected *Sullivan*'s observation that "debate on *public issues* should be uninhibited, robust, and wide-open."[228] But to the extent it failed to apply the actual malice standard, it erred. Rather, Brennan would have held, under the standard proposed by the Court in *Rosenbloom*, that Gertz had to prove actual malice.[229] Public interest, Brennan wrote, may "at times be influenced by the notoriety of the individuals

---

216. *Id.*
217. *Id.*
218. *Gertz*, 418 U.S. at 355-56 (Douglas, J., dissenting).
219. *Id.* at 356 (citing Thomas Jefferson, Drafts of the Kentucky Resolutions of 1798, *in* 8 THE WORKS OF THOMAS JEFFERSON 458, 464-65 (Paul Leicester Ford ed., 1904).
220. *Id.*
221. *Id.* at 356-57.
222. *Id.* at 357 ("With the First Amendment made applicable to the States through the Fourteenth, I do not see how States have any more ability to 'accommodate' freedoms of speech or of the press than does Congress.").
223. *Id.* at 359.
224. *Gertz*, 418 U.S. at 359-60 (Douglas, J., dissenting).
225. *Id.* at 360.
226. *Id.* at 361 (Brennan, J., dissenting) (quoting NAACP v. Button, 371 U.S. 415, 433 (1963)).
227. *Id.* at 361-62.
228. *Id.* (quoting N.Y. Times Co. v. Sullivan, 376 U.S. 254, 270 (1964)).
229. *Id.* at 361

involved."[230] At other times, the person involved will have little, if anything, to do with the public's interest in the underlying event.

Although the Court recognized this, it rejected providing "the same level of constitutional protection that has been afforded the media in the context of defamation of public persons."[231] It did so based on the private individual's lack of access to the media to correct the record and that such individuals had not assumed the risks involved with private life.[232] Brennan rejected these distinctions. *Sullivan* did not posit the actual malice rule because public officials had "'any less interest in protecting [their] reputation than an individual in private life.'"[233] Some public officials had very little, if any, access to media channels above that of a private individual.[234] Additionally, that public officials may have assumed the risk of a defamation charge by entering public service "'bears little relationship either to the values protected by the First Amendment or to the nature of our society.'"[235] Social life, Brennan said, "exposes all of us to some degree of public view," and "'[v]oluntarily or not, we are all "public" men to some degree.'"[236]

Instead of breathing space, Brennan wrote, the Court's holding would promote self-censorship. A negligence standard in private figure libel cases would provide little guidance for the media, leaving them "carefully to weigh a myriad of uncertain factors before publication."[237] Negligence in the context of the news media was a rudderless concept and would leave them to guess "how a jury might assess the reasonableness of steps taken by it to verify the accuracy" of a report's representations.[238] Worse yet, juries that are not sympathetic to the news media or to the politics of a particular report may use the negligence standard to exact damages based on the content of the speech rather than the conduct of the publisher.[239]

Sullivan avoided all these problems, and, Brennan wrote, the majority's doubt in requiring judges to decide whether issues were *public* issues was misplaced.[240] While the task may not "be easy," it did not ask judges to perform any duty that was outside of "their traditional functions."[241] Judges had already applied *Rosenbloom* without difficulty and, similarly, undertaken the public figure analysis in *Curtis Publishing Co.*, both of which required tackling the question of whether something was a public issue.[242] That the "public interest was necessarily broad," alleviated the chances of ambiguous

230. *Gertz*, 418 U.S. at 362.
231. *Id.*
232. *Id.* at 362-63.
233. *Id.* (quoting Rosenbloom v. Metromedia, Inc., 403 U.S. 29, 46-47 (1971) (plurality opinion), *abrogated by Gertz*, 418 U.S. 323).
234. *Id.*
235. *Id.* at 364 (Brennan, J., dissenting) (quoting *Rosenbloom*, 403 U.S. at 47 (plurality opinion)).
236. *Gertz*, 418 U.S. at 364 (quoting *Rosenbloom*, 403 U.S. at 48 (plurality opinion)).
237. *Id.* at 365-66.
238. *Id.* at 366.
239. *Id.* at 367.
240. *Id.* at 368-69.
241. *Id.* at 369
242. *Gertz*, 418 U.S. at 369. (Brennan, J., dissenting).

line drawing both for judges and for the news media trying to assess potential liability.[243] Because Gertz failed to show actual malice, Brennan would have affirmed the decision below.[244]

Last came White's dissenting opinion drawing into question the Court's extension of *Sullivan*.[245] For two hundred years, he began, "the law of defamation and right of the ordinary citizen to recover for false publication injurious to his reputation have been almost exclusively the business of state courts and legislatures."[246] Traditional rules were lenient: a private citizen need only prove a false and defamatory publication, and general damages were presumed.[247] This law had "remained untouched by the First Amendment" until the Court's opinion in *Sullivan*.[248] As White saw it, *Gertz* was an unfortunate extension of that case.

By requiring a showing of some level of fault and limiting damages even in cases related to private figures, the Court had just "federalized major aspects of libel laws."[249] In doing so, it held "unconstitutional in important respects the prevailing defamation law in all or most of the 50 States."[250] While White did not believe the decision was "illegitimate or beyond the bounds of judicial review," he did believe it was "an ill-considered exercise of the power entrusted to this Court," and he worried about the "wholesale" "scuttling" of state libel law and the Court's "deprecating the reputation interest of ordinary citizens."[251]

White split much of the substance of his dissent into two parts. First, he focused on the state of the common law of libel before *Sullivan*. When the Restatement of Torts was published in 1938, it represented the accepted view that "publication in written form of defamatory material . . . subjected the publisher to liability although no special harm to reputation was actually proved."[252] The exceptions were limited to truth being a defense and some statements being privileged.[253] But, "[a]t the very least," these rules "allowed the recovery of nominal damages for any defamatory publication actionable *per se* and thus performed 'a vindicatory function by enabling the plaintiff publicly to brand the defamatory publication as false.'"[254]

Once liability was shown, damages owed for libel or slander *per se* were either the harm to the reputation as shown by a plaintiff, or when a plaintiff failed to make such a showing, the harm that one could expect from

243. *Id.*
244. *Id.*
245. *Id.* at 369, 377 (White, J., dissenting). Burger also dissented, largely without substance. *Id.* at 354 (Burger, C.J., dissenting).
246. *Id.* at 369-70, 377 (White, J., dissenting).
247. *Gertz*, 418 U.S. at 370 (White, J., dissenting).
248. *Id.* at 370.
249. *Id.*
250. *Id.*
251. *Id.*
252. *Id.* at 371 (citing RESTATEMENT OF TORTS § 569 (AM. L. INST. 1938)).
253. *Gertz*, 418 U.S. at 371 (White, J., dissenting) (quoting RESTATEMENT OF TORTS § 569 (AM. L. INST. 1938)).
254. *Id.* at 372 (quoting RESTATEMENT OF TORTS § 569 cmt. b (AM. L. INST. 1938)).

such a defamatory charge.[255] These general damages for loss of reputation were "the heart of the libel-and-slander-per-se damage scheme."[256] They existed because, at least when it came to cases of *per se* defamation, the law assumed "the content of the publication itself was so likely to cause injury."[257] *Gertz*, however, marked a drastic departure from this system by prohibiting a plaintiff from "rest[ing] his case with proof of a libel defamatory on its face."[258]

White believed that these "radical changes in the law" and the "severe invasions of the prerogatives of the States" should "at least be shown to be required by the First Amendment or necessitated by our present circumstances."[259] But the majority showed neither. *Sullivan* and its progeny had "worked major changes in defamation law," but neither "foreclose[d] in all circumstances recovery by the ordinary citizen on traditional standards of liability, and until today, a majority of the Court had not supported the proposition that, given liability, a court or jury may not award general damages in a reasonable amount without further proof of injury."[260]

In the second half of his dissent, White addressed the question of whether the First Amendment required the result in *Gertz*. He began by stating that there was no historical support that the First Amendment limited libel actions in the District of Columbia or U.S. territories.[261] Moreover, "10 of the 14 States that had ratified the Constitution by 1792 had themselves provided constitutional guarantees for free expression, and 13 of the 14 nevertheless provided for the prosecution of [criminal] libels."[262] Before the Revolution, the common law of libel was adopted in the Colonies.[263] Far from a free press being embraced in early America, he said it was "sharply curtailed."[264]

Based on this, White found that there was "[s]cant, if any, evidence . . . that the First Amendment was intended to abolish the common law of libel, at least to the extent of depriving ordinary citizens of meaningful redress against their defamers."[265] Instead, "the common-law rules that subjected the libeler to responsibility for the private injury" were not "abolished by the protection extended to the press in our constitutions."[266] In fact, the Founders,

255. *Id.*
256. *Id.* at 372-73.
257. *Id.* at 373.
258. *Id.* at 375.
259. *Gertz*, 418 U.S. at 376-77 (White, J., dissenting)
260. *Id.*
261. *See id.* at 380.
262. *Id.*
263. *Id.* at 381 (citing Jerome Lawrence Merin, *Libel and the Supreme Court*, 11 WM. & MARY L. REV. 371, 373 (1969)).
264. *Id.*
265. *Gertz*, 418 U.S. at 381 (White, J., dissenting).
266. *Id.* (quoting 2 T.M. COOLEY & WALTER CARRINGTON, A TREATISE ON THE CONSTITUTIONAL LIMITATIONS WHICH REST UPON THE LEGISLATIVE POWER OF THE STATES OF THE AMERICAN UNION 883 (8th ed. 1927)).

he said, viewed freedom of press as meaning only freedom from prior censorship.[267] These views reflected modern scholars', he added.[268]

White also weaponized the ambiguity of the historical record around the Bill of Rights, asserting that the Bill of Rights was "unclear and inconclusive on any articulated intention of the Framers as to the free press guarantee."[269] At best, "Benjamin Franklin, John Adams, and William Cushing favored limiting freedom of the press to truthful statements, while others such as James Wilson suggested a restatement of the Blackstone standard."[270] Jefferson endorsed James Madison's proposed clause protecting the freedom of speech, but offered instead that the "people shall not be deprived of their right to speak, to write, or *otherwise* to publish anything but false facts affecting injuriously the life, liberty, or reputation of others."[271]

Moreover, the Court had recently reiterated the view "that defamatory utterances were wholly unprotected by the First Amendment."[272] In *Near v. Minnesota ex rel. Olson*, the Court wrote "that the common law rules that subject the libeler to responsibility for the public offense, as well as for the private injury, are not abolished by the protection extended in our constitutions."[273] And, in *Chaplinsky v. New Hampshire*, the Court declared, that libelous speech was one of the "well-defined and narrowly limited classes of speech, the prevention and punishment of which have never been thought to raise any Constitutional problem."[274]

The *Sullivan* Court, however, "could not accept the generality of this historic view," finding that "the First Amendment was intended to forbid actions for seditious libel and that defamation actions by public officials were therefore not subject to the traditional law of libel."[275] *Sullivan*, White argued, "reflected one side of the dispute that raged at the turn of the nineteenth century [over the Sedition Act] and also mirrored the views of some later scholars."[276] White then made his dispute with *Gertz* plain, while endorsing *Sullivan*: "[t]he central meaning of [*Sullivan*], *and for me* the First Amendment as it relates to libel laws, is that seditious libel—criticism of government and public officials—falls beyond the police power of the

---

267. *Id.* at 381 (citing LEONARD W. LEVY, LEGACY OF SUPPRESSION: FREEDOM OF SPEECH AND PRESS IN EARLY AMERICAN HISTORY 247-48 (1960)); Jerome Lawrence Merin, *Libel and the Supreme Court*, 11 WM. & MARY L. REV. 371, 376 (1969); John E. Hallen, Comment, *Fair Comment*, 8 TEX. L. REV. 41, 56 (1929); 4 WILLIAM BLACKSTONE, COMMENTARIES *150-53).

268. *Gertz*, 418 U.S. at 381 (White, J., dissenting) (citing ALEXANDER MEIKLEJOHN, POLITICAL FREEDOM: THE CONSTITUTIONAL POWERS OF THE PEOPLE 21 (1965); Robert A. Leflar, *The Free-ness of Free Speech*, 15 VAND. L. REV. 1073, 1080-81 (1962)).

269. *Id.* at 383.

270. *Id.* (citing Jerome Lawrence Merin, *Libel and the Supreme Court*, 11 WM. & MARY L. REV. 371, 377 (1969)).

271. *Id.* at 384 (White, J., dissenting) (quoting FRANK LUTHER MOTT, JEFFERSON & THE PRESS 14 (1943)).

272. *Id.* at 384-85.

273. *Id.* at 385 (quoting Near v. Minnesota *ex rel.* Olson, 283 U.S. 697, 714 (1931)).

274. *Gertz*, 418 U.S. at 385 (White, J., dissenting) (quoting Chaplinsky v. New Hampshire, 315 U.S. 568, 571-72 (1942)).

275. *Id.* at 386.

276. *Id.* at 386-87 (citing LEONARD W. LEVY, LEGACY OF SUPPRESSION: FREEDOM OF SPEECH AND PRESS IN EARLY AMERICAN HISTORY 247-48 (1960)).

State."[277] But, White said, neither *Sullivan* "nor its progeny suggest that the First Amendment intended in all circumstances to deprive the private citizen of his historic recourse to redress published falsehoods damaging to reputation or that, contrary to history and precedent, the Amendment should now be so interpreted."[278]

Philadelphia Newspapers, Inc. v. Hepps. In 1986, the Court considered whether private figure libel plaintiffs had to plead and prove falsity in order to recover.[279] There, the *Philadelphia Inquirer* had published several articles suggesting that a chain of stores had ties to the mob and power to influence government officials and proceedings.[280] Maurice Hepps, the owner of the chain and a private figure, sued, alleging that the articles defamed him.[281] The trial court found that Pennsylvania's statutory scheme, which placed the burden of proving the truth of the disputed statements on the defendant, violated the First Amendment.[282] Therefore, the burden to prove falsity lay with the plaintiff.[283] During trial, the trial judge declined to grant a requested jury instruction that the jury could infer a negative inference from the appellants' failure to disclose sources, and the jury subsequently found for the *Inquirer*.[284] The Pennsylvania Supreme Court reversed, concluding that *Gertz* "simply requir[ed] the plaintiff to show fault," not falsity.[285]

The U.S. Supreme Court disagreed. Despite the plaintiff being a private figure, the Court found that the Constitution required him to show falsity because the case concerned a matter of public interest. As Justice Sandra Day O'Connor explained for the majority, "We believe that the common law's rule on falsity—that the defendant must bear the burden of proving truth—must similarly fall here to a constitutional requirement that the plaintiff bear the burden of showing falsity, as well as fault, before recovering damages."[286] The reason? "[P]lacement by state law of the burden of proving truth upon media defendants who publish speech of public concern deters such speech because of the fear that liability will unjustifiably result."[287] This "chilling effect," O'Connor wrote, was "antithetical to the First Amendment's protection of true speech on matters of public concern."[288]

Milkovich v. Lorain Journal Co. In 1990, the Court considered whether the First Amendment shielded statements of opinion from defamation liability.[289] The underlying dispute related to a local newspaper editorial about a high school wrestling coach, Michael Milkovich, who argued that he was defamed by an implication in the editorial that he perjured

277. *Id.* (emphasis added).
278. *Id.*
279. Phila. Newspapers, Inc. v. Hepps, 475 U.S. 767, 769 (1986).
280. *Id*.
281. *Id.* at 770.
282. *Id.*
283. *Id.*
284. *Id.* at 770-71.
285. *Hepps*, 475 U.S. at 771.
286. *Id.* at 776.
287. *Id.* at 777.
288. *Id.*
289. Milkovich v. Lorain J. Co., 497 U.S. 1, 17 (1990).

himself.[290] After a protracted legal battle, the Ohio Supreme Court, in a related case, found that the challenged defamatory statement was a matter of opinion.[291] As a result, the Ohio court of appeals affirmed judgment for the defendants, and Milkovich sought review by the Supreme Court.[292]

The Court granted review "to consider the important questions raised by the Ohio courts' recognition of a constitutionally required 'opinion' exception to the application of its defamation laws."[293] While the Court declined to adopt some of the broader interpretations of the opinion doctrine developed below in that case (and later reaffirmed on independent state law grounds), it emphasized its holding in *Hepps*: that a defamation plaintiff could only recover if he or she carried his or her burden of proving that the allegedly defamatory statement was a false statement of fact.[294] According to the Court, "we think *Hepps* stands for the proposition that a statement on matters of public concern must be provable as false before there can be liability under state defamation law."[295]

Masson v. New Yorker Magazine. At issue in *Masson* was an article written by Janet Malcolm about a schism among intellectuals at the Sigmund Freud Archives.[296] Malcom interviewed one professor, Jeffrey Masson, who had had a falling out with the Archives and a fact-checker followed up with Masson after Malcom prepared the article.[297] According to Masson, he expressed shock at several errors in the article and, specifically, had questions about certain quotations that Malcom attributed to him.[298] After the *New Yorker* published Malcolm's article and after she later flipped the article into a book, Masson sued for libel, alleging that the misquotations suggested he was a sex-crazed academic.[299]

After the district court granted summary judgment finding the statements to be substantially true and the Ninth Circuit affirmed, the Supreme Court reversed. It found that "[m]inor inaccuracies do not amount to falsity so long as 'the substance, the gist, the sting, of the libelous charge be justified.'"[300] Put differently, "the statement is not considered false unless it 'would have a different effect on the mind of the reader from that which the pleaded truth would have produced.'"[301] For that reason, even the deliberate falsification of words in a quotation would not result in a finding of falsity "unless the alteration results in a material change in the meaning conveyed by the statement."[302]

290. *Id.* at 7.
291. *Id.* at 8-9.
292. *Id.* at 9-10.
293. *Id.* at 10.
294. *Id.* at 19.
295. *Milkovich*, 497 U.S. at 19.
296. *See generally* Masson v. New Yorker Mag., Inc., 501 U.S. 496 (1991).
297. *Id.* at 501.
298. *Id.*
299. *Id.* at 502-03.
300. *Id.* at 517.
301. *Id.*
302. *Masson*, 501 U.S. at 517.

The nearly three decades of precedent, from *Sullivan* to *Masson*, establishes firm, constitutional protections for defendants in defamation cases that allow speakers to pursue topics of public concern by lessening the chilling effect of future libel lawsuits. While the States are free to provide additional safeguards through their constitutions, statutes, or common law, the Supreme Court has established minimum requirements that plaintiffs must satisfy before succeeding in defamation lawsuits under the First and Fourteenth Amendments. Nevertheless, we see in this history disagreements in how best to address the conflict between the law of libel and the First Amendment. Some Justices focused on fault, others on other elements of the claim like the "of and concerning" inquiry and falsity, and still others on limitations of damages. As will be shown, our proposal takes the best of these ideas across majority, concurring, and dissenting opinions to address the weaponization of libel lawsuits and continuing threats to *Sullivan.*

IV. THE WEAPONIZATION OF LIBEL LAWSUITS AND THE DRUMBEAT OF THREATS TO *SULLIVAN*

Ten years ago, the law of libel was a sleepy area of the law—not so much today. While judgments about causation are difficult, the recent glut of libel lawsuits filed against news organizations picked up as former President Donald Trump ran for office in 2016. Trump had long resorted to libel lawsuits and threats, including suing an author whose book he said defamed him by describing him as a millionaire rather than a billionaire.[303] While on the campaign trail in February 2016, Trump announced that he would "open up our libel laws."[304] He added, "So when *The New York Times* writes a hit piece which is a total disgrace or when *The Washington Post*, which is there for other reasons, writes a hit piece, we can sue them and win money instead of having no chance of winning because they're totally protected."[305]

Trump renewed these calls on the eve of the release of Bob Woodward's book *Fear: Trump in the White House*, complaining that "someone can write an article or book, totally make up stories and form a picture of a person that is literally the exact opposite of the fact, and get away

303. Eriq Gardner, *Donald Trump Loses Libel Lawsuit over Being Called a 'Millionaire*,*'* HOLLYWOOD REP. (Sept. 8, 2011, 2:23 PM), https://www.hollywoodreporter.com/business/business-news/donald-trump-loses-libel-lawsuit-232923 [https://perma.cc/2AZ6-TXWT].

304. Daniel Politi, *Donald Trump Vows to Curb Press Freedom Through Harsher Libel Laws*, SLATE (Feb. 27, 2016, 10:24 AM), https://slate.com/news-and-politics/2016/02/donald-trump-vows-to-curb-press-freedom-through-libel-laws.html [https://perma.cc/P3W2-FRXR].

305. Hadas Gold, *Donald Trump: We're Going to 'Open up' Libel Laws*, POLITICO (Feb. 26, 2016, 2:31 PM), https://www.politico.com/blogs/on-media/2016/02/donald-trump-libel-laws-219866 [https://perma.cc/5TXE-69BD].

with it without retribution or cost."[306] Then, in an interview with the *Times*, he said, "We are going to take a strong look at our country's libel laws so that when somebody says something false and defamatory about someone, that person will have meaningful recourse in our courts."[307] Current libel law, he added, was "a sham and a disgrace and do not represent American values or American fairness."[308]

Many were quick to point out that the President has no power to change the law of libel, which is first a feature of state law.[309] But his comments appear to have politicized and publicized the law of libel. Throughout his presidency, many, including his own campaign, increasingly resorted to defamation lawsuits and threats. By early 2020, the Trump campaign filed four lawsuits against *The New York Times*, *The Washington Post*, CNN, and an unlucky local Wisconsin television station that ran a political ad attacking Trump's coronavirus response.[310] As Neal Katyal and Joshua Geltzer observed in *The Atlantic* after the campaign sued the three national news organizations but before they turned their eye on Northern Wisconsin's WJFW-TV, "[E]ven if these lawsuits are unlikely to succeed, they can nevertheless do great harm" through self-censorship, especially by "local media outlets—whether newspapers, radio stations, TV news programs, or websites—that already are struggling to stay afloat."[311]

Devin Nunes, the former Congressman, has filed defamation lawsuit after defamation lawsuit against his critics, including the Rachel Maddow Show, *The Washington Post*, Twitter, CNN, *Esquire Magazine*, and a fake cow's Twitter account.[312] Among other things, these complaints alleged that defendants had "impugn[ed] [Nunes'] reputation and undermine[d] his

306. Matthew Weaver & Joanna Walters, *Trump Dismisses Bob Woodward's Book: 'Lies and Phony Sources*,*'* GUARDIAN (Sept. 5, 2018, 10:18 AM), https://www.theguardian.com/us-news/2018/sep/05/donald-trump-dismisses-bob-woodward-book-fear [https://perma.cc/SK3B-Y3GV].

307. Michael M. Grynbaum, *Trump Renews Pledge to 'Take a Strong Look' at Libel Laws*, N.Y. TIMES (Jan. 10, 2018), https://www.nytimes.com/2018/01/10/business/media/trump-libel-laws.html [https://perma.cc/86AY-YTHU].

308. *Id.*

309. *Id.*

310. Joshua A. Geltzer & Neal K. Katyal, *The True Danger of the Trump Campaign's Defamation Lawsuits*, ATLANTIC (Mar. 11, 2020), https://www.theatlantic.com/ideas/archive/2020/03/true-danger-trump-campaigns-libel-lawsuits/607753 [https://perma.cc/2J6X-FRPK]; Matt Shuham, *Trump Campaign Sues Small Wisconsin TV Station over Critical Super PAC Ad*, TALKING POINTS MEMO (Apr. 13, 2020, 2:45 PM), https://talkingpointsmemo.com/news/trump-campaign-sues-small-wisconsin-tv-station-over-critical-super-pac-ad [https://perma.cc/7AFG-AZXS].

311. Geltzer & Katyal, *supra* note 310.

312. Katie Irby, *Devin Nunes Sues Washington Post. It's His 7th Lawsuit in 12 Months*, FRESNO BEE (Mar. 2, 2020), https://www.fresnobee.com/news/local/article240797241.html [https://perma.cc/WCR3-JRLA].

relationship with the president."[313] Joe Arapaio, the former Maricopa County Sheriff, sued CNN, *Huffington Post*, and *Rolling Stone*, alleging that inaccurate reporting ruined his chances at a 2020 run for Senate.[314] And before that, he sued the *Times* for the same reasons.[315] At that time, his lawyer called Michelle Cottle, a *Times* reporter individually named, a "hate-filled reporter" who worked for a "venomous leftist publication."[316]

In 2017, after the publication of an editorial that some read as implying that Sarah Palin motivated the assassination attempt on Gabby Giffords, Palin sued the *Times*.[317] Palin argued that the editorial could be read as referring to her (although it did not name her) and further that it defamed her by implying that she had motivated the shooter (she had released a map with stylized cross-hairs over congressional districts).[318] After the *Times* won a motion to dismiss, the Second Circuit reversed, allowing the case to go into discovery.[319] At trial, both the judge and the jury sided with the *Times*.[320] The result came after Palin's testimony seemed less focused on the editorial at issue and more on

313. *Id.*; *see also* Dominick Mastrangelo, *Nunes Sues MSNBC, Alleging Rachel Maddow Defamed Him*, HILL (Aug. 4, 2021, 8:30 AM), https://thehill.com/homenews/media/566250-nunes-sues-msnbc-alleging-rachel-maddow-defamed-him [https://perma.cc/Y5KW-6JVP]; Kate Irby, *Devin Nunes' Lawyer Facing Prospect of Sanctions After Two Recent, Rare Court Warnings*, FRESNO BEE (May 4, 2020), https://www.fresnobee.com/news/local/article241064301.html [https://perma.cc/7RQ6-3PWT].

314. Eriq Gardner, *CNN, HuffPost and Rolling Stone Beat Joe Arpaio Libel Lawsuit*, HOLLYWOOD REP. (Oct. 31, 2019, 1:13 PM), https://www.hollywoodreporter.com/thr-esq/cnn-huffpo-rolling-stone-beat-joe-arpaio-libel-lawsuit-1251504 [https://perma.cc/J2K9-FRPE].

315. Jamie Ross, *Arpaio Slaps NY Times with Libel Suit over Op-Ed*, COURTHOUSE NEWS SERV. (Oct. 17, 2018), https://www.courthousenews.com/arpaio-slaps-ny-times-with-libel-suit-over-op-ed [https://perma.cc/VJ2G-XBVZ].

316. *Id.*

317. *See, e.g.*, Marc Tracy, *Sarah Palin's Defamation Suit Against New York Times Is Reinstated*, N.Y. TIMES (Aug. 6, 2019), https://www.nytimes.com/2019/08/06/business/sarah-palin-lawsuit-new-york-times.html [https://perma.cc/RRC5-785E].

318. Oliver Darcy, *Why the* Sarah Palin v. New York Times *Trial Will Be an 'Excruciating Experience' for the Paper*, CNN (Feb. 8, 2022, 11:27 AM), https://www.cnn.com/2022/01/22/media/sarah-palin-new-york-times-trial/index.html [https://perma.cc/F4LL-WS92].

319. Palin v. New York Times Co., 940 F.3d 804 (2d Cir. 2019).

320. Caleb Pershan, *The New York Times Wins Case Against Sarah Palin. Twice.*, COLUM. JOURNALISM REV. (Feb. 15, 2022), https://www.cjr.org/analysis/the-new-york-times-wins-case-against-sarah-palin-twice.php [https://perma.cc/KB6N-SDLG].

general grievances against the *Times* about the "lies" it published about her.[321] We could go on.[322]

Sullivan has also been targeted out of court. In 2022, it was revealed that Florida Governor Ron DeSantis' office sought to pass a bill that would have made it easier to bring defamation cases. As the *Orlando Sentinel* reported, the bill would have challenged "decades-old First Amendment protections for the news media and [made] it easier for high-profile people to win defamation lawsuits."[323] Its goal, a briefing document said, was "to end federal standards established in the *Times* ruling and make defamation purely a matter of state law."[324] Also in 2022, Kyle Rittenhouse, who became a far-right media darling after he was acquitted on charges relating to the deaths of two people in Wisconsin during unrest in 2020, said he would begin selling a video game to raise "funds to sue the left-wing media organizations for defamation."[325]

The resort to libel lawsuits is not only coming from the right. One prominent example is the decade-long battle by climate scientist Michael Mann against the conservative *National Review* and the Competitive Enterprise Institute, among others.[326] While the lawsuit is technically about a criticism of the bona fides of Mann's data, it has transformed into something of a Scopes Trial for climate change. As the *National Review* wrote of the

---

321. David Folkenflik, *Sarah Palin Testifies She Felt Powerless to Fight 'New York Times' over Editorial*, NPR (Feb. 10, 2022, 7:00 PM), https://www.npr.org/2022/02/10/1079861851/nyt-sarah-palin-testifies-defamation-case [https://perma.cc/H43Z-LTFM].

322. *See generally, e.g.*, Ted Johnson, *Judge Tosses Out Project Veritas' Defamation Lawsuit Against CNN*, DEADLINE (Mar. 18, 2022, 9:31 AM) https://deadline.com/2022/03/cnn-project-veritas-defamation-lawsuit-1234981852 [https://perma.cc/5V4R-R7BN]; Ted Johnson, *Judge Sides with ABC, CBS, New York Times and Other Outlets in Libel Cases Filed by Nick Sandmann over Lincoln Memorial Incident*, DEADLINE (July 27, 2022, 11:47 AM), https://deadline.com/2022/07/nick-sandmann-loses-libel-suit-new-york-times-1235079085 [https://perma.cc/Q9UC-QJ5D]; Eriq Gardner, *Mike Lindell Experiences Rejection in Libel Suit over Hollywood Romance*, HOLLYWOOD REP. (Dec. 10, 2021, 1:14 PM), https://www.hollywoodreporter.com/news/general-news/mike-lindell-libel-suit-1235061017 [https://perma.cc/V6Q4-FEB3]; Michael M. Grynbaum, *Lt. Gov. Justin Fairfax of Virginia Sues CBS for Defamation, Seeking $400 Million*, N.Y. TIMES (Sept. 12, 2019), https://www.nytimes.com/2019/09/12/business/media/justin-fairfax-virginia-cbs-defamation.html [https://perma.cc/YN78-2ZME]; Eriq Gardner, *Gizmodo Beats Jason Miller Defamation Lawsuit over "Abortion Pill" Story*, HOLLYWOOD REP. (Aug. 28, 2019, 5:56 AM), https://www.hollywoodreporter.com/business/business-news/gizmodo-beats-jason-miller-defamation-lawsuit-abortion-pill-story-1235101 [https://perma.cc/2JYP-B4H7].

323. Skyler Swisher, *Desantis' Office Considered a Bill to Target Libel Laws, Records Show*, ORLANDO SENTINEL (May 17, 2022, 6:45 PM), https://www.orlandosentinel.com/politics/os-ne-libel-law-draft-bill-20220517-ujyksk3uzzb5rl5y4y6puik3ha-story.html [https://perma.cc/8JR3-TDWF].

324. *Id.*

325. Mikhail Klimentov, *Kyle Rittenhouse Announces Video Game to Fund Media Defamation Suits*, WASH. POST (June 23, 2022, 2:20 PM), https://www.washingtonpost.com/video-games/2022/06/23/kyle-rittenhouse-video-game-defamation [https://perma.cc/74MC-ZX8Y].

326. Marianne Lavelle, *Nine Years After Filing a Lawsuit, Climate Scientist Michael Mann Wants a Court to Affirm the Truth of His Science*, INSIDE CLIMATE NEWS (Feb. 7, 2021), https://insideclimatenews.org/news/07022021/michael-mann-defamation-lawsuit-competitive-enterprise-institute-national-review [https://perma.cc/734E-9WDJ].

litigation, "this is not how we should want to settle political or scientific questions in American life."[327] When the lawsuit reached the Supreme Court on an interlocutory basis, the Court refused to hear it but Justice Samuel Alito dissented, writing, "[R]equiring a free speech claimant to undergo a trial after a ruling that may be constitutionally flawed is no small burden. . . . Those prospects may deter the uninhibited expression of views that would contribute to healthy public debate."[328]

In fact, some of the biggest libel judgments may come against conservative media for their reporting and commentary about the 2020 election and the "big lie." Dominion Voting Systems sued Fox News and various principals for $1.6 billion, alleging that its claims that Dominion voting machines were a conduit for election fraud were false and defamatory.[329] Another voting machine company, Smartmatic, also sued Fox News over similar claims for $2.7 billion.[330] Others, including Rudy Giuliani, Sydney Powell, and My Pillow CEO Mike Lindell were also targeted with libel lawsuits as a result of their crusade against and related commentary about non-existent fraud in the 2020 election.[331] Election workers also sued One America News Network for libel after the network alleged they were involved in election fraud—a lawsuit which the network eventually settled.[332]

The rich and powerful, domestic and international, also often sue, hoping to discourage critical speech. Oleg Deripaska, the Russian oligarch, sued the *Associated Press* over reporting he viewed as improperly connecting him to Russian meddling in the 2016 election.[333] Along the same lines, Russian tech entrepreneur Aleksej Gubarev, to which a passing reference was made in the "Steele Dossier," sued *BuzzFeed News* over its publication of the

---

327. *Michael Mann's Lawsuit Stumbles On*, NAT'L REV. (July 30, 2021, 2:53 PM), https://www.nationalreview.com/2021/07/michael-manns-lawsuit-stumbles-on [https://perma.cc/N83F-8VTF].

328. Nat'l Rev., Inc. v. Mann, 140 S. Ct. 344, 348 (2019).

329. Jan Wolfe, *Trial Date Set in Defamation Suit Against Fox News over U.S. Election Claims*, REUTERS (Apr. 13, 2022, 4:51 PM), https://www.reuters.com/world/us/trial-date-set-defamation-suit-against-fox-news-over-us-election-claims-2022-04-12 [https://perma.cc/KQT6-68RG].

330. Jonah E. Bromwich & Ben Smith, *Fox News Is Sued by Election Technology Company for over $2.7 Billion*, N.Y. TIMES (Feb. 4, 2021), https://www.nytimes.com/2021/02/04/business/media/smartmatic-fox-news-lawsuit.html [https://perma.cc/WUY4-CKEH].

331. Ali Swenson, *No Ruling Yet in Dominion Lawsuits Against Powell and Giuliani*, ASSOCIATED PRESS (May 11, 2022), https://apnews.com/article/fact-check-dominion-lawsuits-giuliani-powell-577808419035 [https://perma.cc/AD2U-9DMZ]; Maya Yang, *MyPillow CEO Sued for Defamation by Former Dominion Voting Employee*, GUARDIAN (Apr. 6, 2022, 4:25 PM), https://www.theguardian.com/us-news/2022/apr/06/mypillow-ceo-mike-lindell-sued-defamation-former-dominion-employee [https://perma.cc/9492-HKHX].

332. Jonathan Allen, *Two Atlanta Poll Workers Settle Defamation Lawsuit Against One America*, REUTERS (Apr. 21, 2022, 8:45 PM), https://www.reuters.com/world/us/two-atlanta-poll-workers-settle-defamation-lawsuit-against-one-america-2022-04-21 [https://perma.cc/B8BU-BRS2].

333. Britain Eakin, *Russian Oligarch Loses Defamation Suit Against AP*, COURTHOUSE NEWS SERV. (Oct. 18, 2017), https://www.courthousenews.com/russian-oligarch-loses-defamation-suit-ap [https://perma.cc/582F-CCKH]; *see generally* Deripaska v. Associated Press, 282 F. Supp. 3d 133 (D.D.C. 2017).

same document.[334] Russian oligarchs Mikhail Fridman, Petr Aven, and German Khan also filed suit over the dossier, this time against the intelligence firm Fusion GPS and its founder.[335] That lawsuit became untenable after the Russian Federation waged an illegal war on Ukraine and the international community imposed sanctions on Russian businesses and oligarchs.[336]

Celebrities are repeat defamation parties too, sometimes bringing suits against the media, other times against each other. Johnny Depp and Amber Heard famously sued each other both in England (where Heard won) and in the United States (where Heard lost in large part).[337] A cave diver who rose to prominence after a youth Thai soccer team became trapped in a cave sued Elon Musk after he called him the "pedo guy."[338] Dr. Luke sued Kesha after she made allegations of sexual assault against him.[339]

Often, though, the media was the defendant when it came to celebrity #MeToo allegations. *BuzzFeed News* caught a lawsuit (in Ireland, likely to avoid U.S. law like *Sullivan*) after Tony Robbins, the famed self-help guru, took umbrage at the news outlet's reporting on alleged sexual misconduct.[340] Roy Moore sued Sacha Baron Cohen and Showtime after he appeared in a spoof skit that touched on sexual misconduct allegations against Moore—a spoof that included, according to the Second Circuit, "the obviously farcical pedophile-detecting 'device,' which no reasonable person could believe to be an actual, functioning piece of technology."[341]

Academics sue too. In 2020, Lawrence Lessig, the well-known liberal Harvard professor and former presidential candidate, sued the *Times* over a disagreement as to the import of a blog post he wrote regarding Jeffrey

334. Jaclyn Peiser, *BuzzFeed Wins Defamation Lawsuit Filed by Executive Named in Trump Dossier*, N.Y. TIMES (Dec. 19, 2018), https://www.nytimes.com/2018/12/19/business/media/buzzfeed-dossier-lawsuit-trump-steele-russia.html [https://perma.cc/C86H-NMTK].

335. Eric Tucker, *Oligarchs Drop Defamation Lawsuit over Trump-Russia Dossier*, ABC NEWS (Mar. 18, 2022, 5:17 PM), https://abcnews.go.com/Politics/wireStory/oligarchs-drop-defamation-lawsuit-trump-russia-dossier-83536037 [https://perma.cc/YSU9-KAHK].

336. *Id.*

337. Jeremy W. Peters, *Depp Trial Exposes Risks to Media in Airing #MeToo Accusations*, N.Y. TIMES (June 3, 2022), https://www.nytimes.com/2022/06/03/business/media/depp-heard-journalism-metoo.html [https://perma.cc/F9P2-HEEW].

338. *Elon Musk Wins Defamation Case over 'Pedo Guy' Tweet About Caver*, BBC (Dec. 6, 2019), https://www.bbc.com/news/world-us-canada-50695593 [https://perma.cc/A6K3-G28T].

339. Althea Legaspi, *Kesha Loses in Appeal of Dr. Luke Defamation Ruling*, ROLLING STONE (Apr. 22, 2021), https://www.rollingstone.com/music/music-news/dr-luke-scores-big-win-kesha-defamation-dispute-949226 [https://perma.cc/E92H-E46F].

340. Tamar Lapin, *Tony Robbins Sues Buzzfeed for Defamation over Sexual Misconduct Reports*, N.Y. POST (Nov. 26, 2019, 8:44 PM), https://nypost.com/2019/11/26/tony-robbins-sues-buzzfeed-for-defamation-over-sexual-misconduct-reports [https://perma.cc/6944-M8DU].

341. Moore v. Baron Cohen, No. 21-1702-CV, 2022 WL 2525722, at *3 (2d Cir. July 7, 2022).

Epstein's donations to academic institutions.[342] Carlo Croce, a cancer researcher who has had several articles retracted, sued another academic and the *Times* over statements published by the *Times*.[343] Alan Dershowitz, the Harvard Law School emeritus professor, sued Netflix and CNN, as he lamented being "canceled" after becoming one of Trump's chief legal defenders.[344] Another Harvard Law professor sued *New York Magazine* after it published a devastating profile about how he was apparently conned by two individuals.[345] One New York University professor even sued his colleagues "after they complained to administrators about his encouraging students to question whether masks actually prevent COVID-19 from spreading."[346]

Even criminal libel law is showing a resurgence. In 2022, a federal judge halted an investigation into a political ad by the North Carolina Attorney General initiated by a political opponent.[347] That same year, police arrested a critic of a local police department for criminal libel, but the investigation was abandoned and a federal court later let a federal lawsuit brought by the critic go forward.[348] Stories like these are easy to find. In 2019, New Hampshire police arrested a Facebook warrior critical of the police.[349] The same year, a police officer had his ex-wife arrested under Georgia's

342. *See* Jasper G. Goodman, *Harvard Law School Professor Lawrence Lessig Sues the New York Times for Defamation*, HARV. CRIMSON (Jan. 16, 2020), https://www.thecrimson.com/article/2020/1/16/lessig-nyt-lawsuit [https://perma.cc/EU32-8XHA].

343. Richard Van Noorden, *Exclusive: Investigators Found Plagiarism and Data Falsification in Work from Prominent Cancer Lab*, NATURE (July 20, 2022), https://www.nature.com/articles/d41586-022-02002-5 [https://perma.cc/5K6K-UNFC].

344. Winston Cho, *Netflix, Alan Dershowitz Drop Claims over Jeffrey Epstein Docuseries*, HOLLYWOOD REP. (Mar. 23, 2022, 2:02 PM), https://www.hollywoodreporter.com/business/business-news/netflix-alan-dershowitz-drop-claims-over-jeffrey-epstein-docuseries-1235117657 [https://perma.cc/QBR9-QEWA]; Eriq Gardner, *CNN Can't Dodge Alan Dershowitz Libel Suit*, HOLLYWOOD REP. (May 25, 2021, 12:58 PM), https://www.hollywoodreporter.com/tv/tv-news/cnn-cant-dodge-alan-dershowitz-libel-suit-1234958859 [https://perma.cc/R3WH-6JZT].

345. *See* Debra Cassens Weiss, *Harvard Law Prof Claims Reporter Sexually Harassed Him, Twisted Facts in 'Gullible Man' Story*, A.B.A. J. (Aug. 11, 2020, 1:51 PM), https://www.abajournal.com/news/article/harvard-law-prof-claims-reporter-sexually-harassed-him-twisted-facts-in-gullible-man-story [https://perma.cc/U5L6-N629].

346. Noah Manskar, *NYU Professor Sues Colleagues amid COVID-19 Mask Controversy*, N.Y. POST (Dec. 2, 2020, 4:24 PM), https://nypost.com/2020/12/02/nyu-prof-mark-crispin-miller-sues-colleagues-amid-mask-flap [https://perma.cc/H5BQ-VRXM].

347. Elura Nanos, *Federal Judge Blocks Criminal Libel Investigation of N.C. AG's Campaign Ad, Agrees Law Likely Violates First Amendment*, LAW & CRIME (July 26, 2022, 1:15 PM), https://lawandcrime.com/first-amendment/federal-judge-blocks-criminal-libel-investigation-of-n-c-ags-campaign-ad-agrees-law-likely-violates-first-amendment [https://perma.cc/3N7R-6UTN].

348. Eugene Volokh, *Criminal Libel Arrest for Criticism of Police Officer Was Unconstitutional*, REASON (May 14, 2022, 11:30 AM), https://reason.com/volokh/2022/05/14/criminal-libel-arrest-for-criticism-of-police-officer-was-unconstitutional [https://perma.cc/ENB6-6XCZ].

349. Adam Liptak, *He Disparaged the Police on Facebook. So They Arrested Him*, N.Y. TIMES (Jan. 7, 2019), https://www.nytimes.com/2019/01/07/us/politics/criminal-libel-laws-lawsuit.html [https://perma.cc/W7D2-FBDE].

criminal libel law after she criticized his parenting on Facebook.[350] As the journalist covering the case observed, criminal libel laws today "are almost always used by government employees to silence critics."[351] And in 2022, Washington State adopted a new statute that allows judges to issue orders of protection that "effectively criminalize[s] future libels" and acts as a "mini-criminal-libel law."[352]

This is not the first time the United States has found itself "in the midst of a rejuvenation of the law of libel."[353] Recognizing the scope of the problem in the 1980s, Professor Rodney Smolla explained that "defendants span a spectrum of size, wealth, power, and respectability, ranging from the mainstream orthodoxy of the national-news giants, to local news outlets, to the more sensational press."[354] Potential plaintiffs were similarly varied, including President Jimmy Carter who obtained an apology from the *Washington Post* after it published a column suggesting that the Carter Administration had bugged the Blair House, where incoming President Ronald Reagan was staying.[355] Many of these plaintiffs had also "previously profited from media attention" and included people "deeply involved in the political process," as well as entertainers and writers, among others.[356]

Writing in the same decade, Professor Richard Epstein questioned whether *Sullivan* had really solved anything and called the law of libel "more controversial today" than it was during the 1970s.[357] He added, "It is a commonplace observation that the concern, not to say anxiety, about the threat that defamation actions hold out to freedom of speech and the press has grown mightily, especially in the last decade."[358] Had *Sullivan* been right, one would have expected defamation lawsuits to recede. The trend, however, was "the reverse, for without question the law of defamation is far more controversial today than it was a decade ago, even though there has been little significant change in the framework of the substantive law."[359]

Anthony Lewis sensed something afoot too. "Although [the U.S. press] is the freest in the world, and freer now than it ever has been, it often cries

---

350. Tim Cushing, *Georgia Woman Takes Home $100,000 Settlement After Bogus Criminal Defamation Arrest By Her Ex-Husband (and Current Deputy)*, TECHDIRT (Oct. 30, 2019, 3:23 AM), https://www.techdirt.com/2019/10/30/georgia-woman-takes-home-100000-settlement-after-bogus-criminal-defamation-arrest-her-ex-husband-current-deputy/ [https://perma.cc/99C3-LE2R].

351. *Id.*

352. Eugene Volokh, *Criminal Libel Law, Partly Coming Back in Washington State in Harassment Order Cases*, REASON (June 20, 2022, 12:01 PM), https://reason.com/volokh/2022/06/20/criminal-libel-law-partly-coming-back-in-washington-state-in-harassment-order-cases/ [https://perma.cc/AY3J-PNXX].

353. Rodney A. Smolla, *Let the Author Beware: The Rejuvenation of the American Law of Libel*, 132 U. PA. L. REV. 1, 1 (1983).

354. *Id.* at 2.

355. *Id.* at 2-3.

356. *Id.* at 2.

357. Richard A. Epstein, *Was* New York Times v. Sullivan *Wrong?*, 53 U. CHI. L. REV. 782, 783 (1986).

358. *Id.*

359. *Id.*

that doom is at hand," he wrote in 1983.[360] While he was skeptical of claims by the press, he said that he "must admit that there is something to the concern."[361] Libel, he wrote in advocating that the First Amendment should be reinterpreted as banning libel lawsuits brought against "critics of official conduct," "is not the only form of litigation afflicted in this country today by endless discovery, high costs, extravagant jury verdicts."[362] Still, he argued, it did "not follow that the critics of official conduct must wait for general reforms in our law to get relief from burdens that induce self-censorship."[363] Instead, he urged the Court to act to "find a new remedy."[364]

Many of the same concerns felt in the 1980s are felt today. Libel is now, as it was then, one of the most controversial corners of the law. And, just as in the 1980s, *Sullivan* has not stemmed the rising tide of the suits, nor the rising costs of that litigation. As the Media Law Resource Center observed in the most comprehensive report on *Sullivan* to date, "After a slowdown in the late 2000s and early 2010s, there seems to have been a resurgence" in libel lawsuits "in more recent years after the political climate grew hot during the Trump era."[365] This is not to say that *Sullivan* was wrong, although it might raise the question of whether it went far enough. Despite the glut of libel lawsuits (or because of it), Lee Levine, one of the country's preeminent First Amendment lawyers, said that *Sullivan* remained "a 'landmark' decision that has indeed 'shaped our history' and defined us as a nation."[366]

The weaponization of libel lawsuits is particularly concerning amidst the increasing drumbeat of calls for the Supreme Court to revisit *Sullivan*. The demands to reconsider the actual malice standard intensified in February 2019, when Justice Clarence Thomas wrote a concurrence to a certiorari denial in *McKee v. Cosby*, a defamation case. Although Thomas agreed with his colleagues that the Supreme Court should not review the "factbound question" of whether the plaintiff was properly classified as a limited-purpose public figure, he wrote that it "should reconsider the precedents that require courts to ask it in the first place."[367]

Thomas argued that *Sullivan* and its progeny "were policy-driven decisions masquerading as constitutional law" and that the First Amendment, when it was drafted, was not understood to require actual malice in defamation cases.[368] Thomas's concurrence suggested not only that *Sullivan* should be revisited, but that the First Amendment does not provide any protection to libel defendants. "Historical practice further suggests that

360. Anthony Lewis, New York Times v. Sullivan *Reconsidered: Time to Return to 'the Central Meaning of the First Amendment,'* 84 COLUM. L. REV. 603, 603 (1983).

361. *Id.*

362. *Id.* at 621, 624.

363. *Id.* at 624.

364. *Id.*

365. *Introduction* to MEDIA L. RES. CTR., NEW YORK TIMES V. SULLIVAN: THE CASE FOR PRESERVING AN ESSENTIAL PRECEDENT 1, 7 (2022), https://live-medialaw.pantheonsite.io/wp-content/uploads/2022/03/nytsullivanwhitepaper-1.pdf [https://perma.cc/47JV-LELG].

366. *Id.*

367. McKee v. Cosby, 139 S. Ct. 675, 675-76 (2019) (Thomas, J., concurring in denial of certiorari).

368. *Id.* at 676

protections for free speech and a free press—whether embodied in state constitutions, the First Amendment, or the Fourteenth Amendment—did not abrogate the common law of libel," he wrote.[369] Yet even under his originalist framework, Thomas ignored more than a century of common law protections from libel lawsuits.[370]

Two years later, in a dissent from the denial of certiorari in *Berisha v. Lawson*, Thomas once again argued that the Court should revisit *Sullivan*.[371] He focused not only on what he believed was the lack of historical support for the actual malice rule, but also on the modern, practical impacts of constitutional protections for defamation defendants.[372] Among the cases that Thomas cited was an online conspiracy theory in 2016 that alleged Democrats had operated a child sex trafficking ring at a Washington, D.C. pizza restaurant, causing an armed gunman to visit the shop.[373] "Our reconsideration is all the more needed because of the doctrine's real-world effects," Thomas wrote. "Public figure or private, lies impose real harm."[374]

Yet Thomas failed to explain how eliminating the actual malice rule would meaningfully reduce the proliferation of conspiracies such as PizzaGate, which were distributed by scores of often anonymous online bulletin board posters. The subjects of the pizza conspiracy included Hillary Clinton and her campaign chair, John Podesta, and it is questionable whether they would have the interest in suing anonymous online posters and drawing even more attention to their ridiculous claims.[375]

In *Berisha*, Thomas was not alone in his calls to rethink *Sullivan*. Justice Neil Gorsuch questioned whether *Sullivan* has led to less responsible journalism. "It seems that publishing without investigation, fact-checking, or editing has become the optimal legal strategy," Gorsuch wrote. "Under the actual malice regime as it has evolved, 'ignorance is bliss.' Combine this legal incentive with the business incentives fostered by our new media world and the deck seems stacked against those with traditional (and expensive) journalistic standards—and in favor of those who can disseminate the most

369. *Id.* at 681.

370. *See* Matthew L. Schafer, *A Response to Justice Thomas*, *in* NEW YORK TIMES V. SULLIVAN: THE CASE FOR PRESERVING AN ESSENTIAL PRECEDENT 9, 77-78 (2022), https://live-medialaw.pantheonsite.io/wp-content/uploads/2022/03/nytsullivanwhitepaper-1.pdf [https://perma.cc/47JV-LELG] ("On the contrary, history amply supports what the Court did in *Sullivan*. Far from being out of step with history, *Sullivan* is the obvious next step in what was then more than 150 years of tussling between libel and freedom of the press. Republicanism, freedom of the press, actual malice, the role of public officials and public figures – it is all in these dusty pages. It was all there long before L.B. Sullivan sued the *New York Times*.").

371. *See generally* Berisha v. Lawson, 141 S. Ct. 2424 (2021) (Thomas, J., dissenting from denial of certiorari). By way of disclosure, one of the article authors, Matthew Schafer, was in-house counsel to Simon & Schuster in the case.

372. *Id.* at 2425.

373. *Id.*

374. *Id.*

375. Gregor Aisch, Jon Huang, & Cecilia Kang, *Dissecting the #PizzaGate Conspiracy Theories*, N.Y. TIMES (Dec. 10, 2016), https://www.nytimes.com/interactive/2016/12/10/business/media/pizzagate.html [https://perma.cc/ZC2F-M9QD].

sensational information as efficiently as possible without any particular concern for truth."[376] These claims, though, lacked evidentiary basis and failed to establish a connection between *Sullivan* and the lack of rigorous journalism. As Levine wrote of the law review article on which Gorsuch based his dissent, it "reads (to paraphrase then-Justice Rehnquist) 'much like a treatise about cooking by someone who has never cooked before, and has no intention of starting now.'"[377]

Finally, Thomas once again called for the reconsideration of *Sullivan* in a June 2022 denial of certiorari in *Coral Ridge Ministries Media v. Southern Poverty Law Center*.[378] Curiously, while the case was rescheduled for consideration for many weeks, Thomas wrote alone, largely regurgitating prior arguments in a short opinion.[379] No other Justice wrote, raising the question of whether another Justice was writing something that he or she ultimately decided not to publish.

Some lower court judges have echoed the calls of Thomas and Gorsuch, including judges on the Florida Court of Appeals[380] and Michigan Court of Appeals.[381] Among the most vociferous criticisms of *Sullivan* came from D.C. Circuit Judge Laurence Silberman. In a 2021 partial dissent, Silberman urged the Supreme Court to overturn *Sullivan*.[382] Rather than focusing only on an originalist critique or the harms of online conspiracy theories, Silberman linked *Sullivan* with what he viewed as the liberal bias of the media and technology companies. "The First Amendment guarantees a free press to foster a vibrant trade in ideas," Silberman wrote. "But a biased press can distort the marketplace. And when the media has proven its willingness—if not eagerness—to so distort, it is a profound mistake to stand by unjustified legal rules that serve only to enhance the press' power."[383]

While recent criticisms have come mainly from conservative judges, liberals are not entirely happy with *Sullivan* either. In 1993, when she was a law professor at the University of Chicago, Justice Elena Kagan wrote a book review in which she highlighted both successes and weaknesses of the landmark case.[384] "The obvious dark side of the *Sullivan* standard is that it

---

376. *Berisha*, 141 S. Ct. at 2428 (Gorsuch, J., dissenting from denial of certiorari) (internal citations omitted).

377. Lee Levine, *Afterward* to MEDIA L. RES. CTR., NEW YORK TIMES V. SULLIVAN: THE CASE FOR PRESERVING AN ESSENTIAL PRECEDENT 193, 194 (2022) (quoting Anderson v. Liberty Lobby, Inc., 477 U.S. 242, 269 (1986) (Rehnquist, J., dissenting)), https://live-medialaw.pantheonsite.io/wp-content/uploads/2022/03/nytsullivanwhitepaper-1.pdf [https://perma.cc/47JV-LELG].

378. *See generally* Coral Ridge Ministries Media, Inc. v. S. Poverty L. Ctr., 142 S. Ct. 2453 (2022) (Thomas, J., dissenting from denial of certiorari).

379. Schafer, *supra* note 52, at 86.

380. Mastandrea v. Snow, 333 So. 3d 326, 328 (Fla. Dist. Ct. App. 2022) (Thomas, J., concurring).

381. Reighard v. ESPN, Inc., No. 355053, 2022 WL 1513112, at *19 (Mich. Ct. App. May 12, 2022) (Boonstra, P.J., concurring).

382. Tah v. Glob. Witness Publ'g, Inc., 991 F.3d 231, 243 (D.C. Cir.) (Silberman, J., dissenting in part), *cert. denied*, 211 L. Ed. 2d 252, 142 S. Ct. 427 (2021).

383. *Id.* at 256.

384. *See generally* Kagan, *supra* note 11.

allows grievous reputational injury to occur without monetary compensation or without any other effective remedy," Kagan wrote.[385]

Perhaps Kagan has changed her mind nearly three decades later, and none of the other Justices will join the calls of Thomas and Gorsuch to revisit *Sullivan*. But as seen in June 2022, when the Supreme Court overturned *Roe v. Wade*, even the most fundamental constitutional liberties are at risk of being overturned at the whim of five justices who disagree with the precedent.[386] Even if Thomas and Gorsuch do not currently have three other votes to overturn *Sullivan,* there is no guarantee that this will always be the case. Nor is there any guarantee that they will be unable to marshal two more votes to at least force reconsideration of *Sullivan*—even if they are ultimately unsuccessful in overturning it.

Rather than stand by and watch decades of vital First Amendment precedent suddenly disappear one day in June, Congress can take steps now to codify *Sullivan* and its progeny and, where necessary, strengthen them to stem the rising tide of politically-motivated defamation lawsuits. It could do so as a matter of federal statutory law by preempting state laws that are inconsistent with the principles laid out in those decisions. As such, we next discuss preemption law as it relates to defamation and then propose our statutory language to address threats to *Sullivan*.

V. PREEMPTION AND DEFAMATION

Although state common law and statutes govern the substantive standards of defamation litigation, federal statutes could partly or entirely preempt state defamation rules,[387] just as the Supreme Court's interpretation of the First Amendment has shaped the contours of state defamation law over the past half century. In other words, Congress could set the minimum protections for defendants in defamation lawsuits. By doing so, it can insulate the press and the public from wild swings in the law of libel should *Sullivan* be overruled, including, especially, the partisan weaponization of libel to punish disfavored speakers.

We recognize that preemption of state common law is a heavy-handed step that requires precise statutory drafting. As the Supreme Court wrote, "because the States are independent sovereigns in our federal system, we have long presumed that Congress does not cavalierly preempt state-law causes of action."[388] Still, our approach is not without precedent. For years, going back at least to the late nineteenth century, commentators have argued for a

385. *Id.* at 205.

386. Dobbs v. Jackson Women's Health Org., 213 L. Ed. 2d 545, 142 S. Ct. 2228 (2022).

387. *See generally* JAY B. SYKES & NICOLE VANATKO, CONG. RSCH. SERV., R45825, FEDERAL PREEMPTION: A PRIMER (2019).

388. Medtronic, Inc. v. Lohr, 518 U.S. 470, 485 (1996).

national libel law.[389] The idea remains popular today, with some arguing for a national libel law based on the Restatement.[390] Other proposals, including a 1980s "study bill" from then-Congressman Chuck Schumer, have sought more limited reforms by, for example, substituting money damages for declaratory relief.[391] Ultimately, Schumer's bill was left to die on the vine.

Our proposal occupies the middle ground. In making it, we look to somewhat recent history. Indeed, were our proposal adopted, it would not be the first time that Congress has sought to preempt state defamation law. Section 230 of the Communications Decency Act, passed as part of the Telecommunications Act of 1996, has preempted many defamation lawsuits against online service providers by partly preempting the common law libel doctrine of republication.[392] The preemption provision of the Freedom of Speech and Press Act is based on Section 230's preemption section.

Congress passed Section 230 in response to concerns over an interpretation of the common law defamation rules in New York, as applied to commercial online services.[393] In 1991, a New York federal judge granted summary judgment for CompuServe in a libel case, reasoning that like newsstands, bookstores, and other distributors, it was liable only if it knew or had reason to know of the defamatory content.[394] "CompuServe has no more editorial control over such a publication than does a public library, book store [sic], or newsstand, and it would be no more feasible for CompuServe to examine every publication it carries for potentially defamatory statements than it would be for any other distributor to do so," the judge wrote.[395]

But in 1995, a New York state court refused to apply the same "distributor" liability standard to Prodigy in a defamation lawsuit seeking $200 million in damages arising from a user's post on a financial discussion board.[396] Because Prodigy's moderation practices were more extensive than those of CompuServe, the judge ruled, it exercised sufficient "editorial control" to face the same liability for all user content as the subscribers who posted it.[397] "Prodigy's conscious choice, to gain the benefits of editorial control, has opened it up to a greater liability than CompuServe and other computer networks that make no such choice," the judge wrote.[398]

---

389. Walter Williams, President, Nat'l Ed. Assoc., Annual Address Before the Tenth Convention of the National Editorial Association (July 2, 1894), *in* 1 THE FIRST DECENNIUM OF THE NATIONAL EDITORIAL ASSOCIATION OF THE UNITED STATES 534, 544 (B.B. Herbert ed., 1896) ("The movement for a national libel law merits consideration. Such a law, rightly framed, would not only close Federal courts to many vexatious suits, but would form the basis for State legislation of a like character.").

390. Alexandra M. Gutierrez, Comment, *The Case for a Federal Defamation Regime*, 131 YALE L.J. F. 19, 44-47 (2021) (discussing preemption).

391. H.R. 2846, 99th Cong. (1985).

392. *See* JEFF KOSSEFF, THE TWENTY-SIX WORDS THAT CREATED THE INTERNET, 4 (2019).

393. *Id.*

394. Cubby, Inc. v. CompuServe, Inc., 776 F. Supp. 135, 141 (S.D.N.Y. 1991).

395. *Id.* at 140.

396. Stratton Oakmont, Inc. v. Prodigy Servs. Co., No. 31063/94, 1995 WL 323710, at *2 (N.Y. Sup. Ct. May 24, 1995).

397. *Id.* at *13.

398. *Id.*

The two New York rulings meant that online service providers could reduce their potential liability for user content by taking a hands-off approach to moderation. This was of particular concern in 1995, as internet connections began to proliferate in homes, schools, and libraries, and legislators and media outlets panicked over the possibility of children accessing pornography on computers.[399] Why have a rule that discourages online services from blocking inappropriate content?

Two congressmen quickly came up with a solution. Within weeks of the ruling against Prodigy, Chris Cox and Ron Wyden introduced the Internet Freedom and Family Empowerment Act, which would later be known as Section 230.[400] During the brief floor discussion of the proposal in 1995, Cox said that "the existing legal system provides a massive disincentive for the people who might best help us control the Internet to do so."[401] Much of the discussion on the House floor focused on the need for companies to provide users with tools to block harmful content and the dangers of the government stepping in to censor. As introduced, Section 230(d) of the bill stated that the Federal Communications Commission has no authority "with respect to economic or content regulation of the Internet or other interactive computer services."[402]

Section 230(c)(1), which received little discussion at the time, states: "No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider."[403] These twenty-six words eliminate the quirk in the common law that caused the New York court to classify Prodigy as a "publisher" because it exercised too much "editorial control." Under Section 230, a platform is not treated as a publisher of third-party content regardless of whether and how it moderates content.[404]

As Cox and Wyden first introduced the bill, it did not address the extent to which it preempts state law. In August 1995, the House attached Section 230 to its version of a massive overhaul of U.S. telecommunications law.[405] In its version of the telecommunications bill, the Senate tried to address minors' access to online pornography in a very different way: its Communications Decency Act imposed criminal penalties for the transmission of indecent material.[406]

---

399. *See* KOSSEFF, *supra* note 401, at 61-62.

400. *Id.* at 64.

401. 104 CONG. REC. H8469 (daily ed. Aug. 4, 1995) (statement of Rep. Christopher Cox).

402. H.R. 1978 (104th Cong.).

403. 47 U.S.C. § 230(c)(1). The statute also prevents interactive computer service providers and users from being liable for "any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected," and providing the technical means to do so. 47 U.S.C. § 230(c)(2).

404. Zeran v. Am. Online, Inc., 129 F.3d 327, 330 (4th Cir. 1997) ("Thus, lawsuits seeking to hold a service provider liable for its exercise of a publisher's traditional editorial functions – such as deciding whether to publish, withdraw, postpone or alter content – are barred.").

405. *See* Kosseff, *supra* note 401 at 70.

406. *Id.* at 57-78.

In the conference committee, both Section 230 and the Communications Decency Act were merged into the same section of the final telecommunications law.[407] But Section 230 underwent some last-minute changes in the conference committee. Most of the changes were minor, but the conferees deleted the restrictions on the FCC's authority (perhaps to avoid conflict with the Senate's indecency provisions).[408] The conferees added another sentence, in Section 230(e)(3), that has proven to be key to preemption of defamation and other state claims: "No cause of action may be brought and no liability may be imposed under any State or local law that is inconsistent with this section."[409] That sentence was added directly after a line that had been in the earlier version: "Nothing in this section shall be construed to prevent any State from enforcing any State law that is consistent with this section."[410] The deletion of the FCC provision and the addition of the preemption language made clear that Section 230 was not merely about anti-regulation, but that it was intended to limit litigation against platforms arising from user-generated content.

The scope of Section 230's preemptive effect became clear throughout 1996 and 1997, as *Zeran v. America Online* was litigated. *Zeran* arose from hoax AOL bulletin board posts from an anonymous user, purporting to sell t-shirts with crude jokes about the recent Oklahoma City bombing.[411] The posts included the plaintiff's first name and phone number. Despite the plaintiff's repeated calls to AOL to inform them that he had nothing to do with the posts, the company failed to prevent additional posts.[412] The plaintiff sued AOL for negligently distributing defamatory posts.[413]

Zeran was the first federal district and appellate court interpretation of Section 230 and is best known for broadly interpreting Section 230(c)(1) to preclude not only publisher liability but also distributor liability (which is imposed if the defendant knows or has reason to know of the content at issue). In other words, even if an online platform receives a complaint about defamatory or otherwise harmful user content and fails to remove it, the platform still is not liable for that content.[414]

The other important—though less obvious—holding of *Zeran* is the preemptive effect of Section 230 on state laws.[415] When District Court Judge T.S. Ellis granted summary judgment for AOL in March 1997, he engaged in an extensive analysis that ultimately concluded that Section 230(e)(3) preempted state tort claims, including negligence.

---

407. Jeff Kosseff, *What's in a Name? Quite a Bit, if You're Talking About Section 230*, LAWFARE (Dec. 19, 2019, 1:28 PM), https://www.lawfareblog.com/whats-name-quite-bit-if-youre-talking-about-section-230 [https://perma.cc/RTN5-GLWD].

408. *Id.*

409. 47 U.S.C. § 230(e)(3).

410. *Id.*

411. *Zeran*, 129 F.3d at 329.

412. *Id.*

413. *Id.*

414. *Id.* at 330.

415. *Id.* at 334.

Preemption takes two forms: express or implied.[416] Ellis first determined that Section 230 did not expressly preempt all state tort claims. To arrive at that conclusion, Ellis pointed to the express preemption provision of the Employee Retirement Income Security Act, which states that "provisions of this subchapter and subchapter III of this chapter shall supersede any and all State laws insofar as they may now or hereafter relate to any employee benefit plan described in section 1003(a) of this title and not exempt under section 1003(b) of this title," and defines "state law" as "all laws, decision, rules, regulations, or other State action having the effect of law, of any State."[417] ERISA's preemption provision excludes certain categories of state laws, such as banking, which Ellis took to mean that ERISA "explicitly defines the extent to which Congress intended federal preemption of state law."[418]

In contrast, all of Section 230(e)(3), Ellis noted, states: "Nothing in this section shall be construed to prevent any State from enforcing any State law that is consistent with this section. No cause of action may be brought and no liability may be imposed under any State or local law that is inconsistent with this section."[419] The two sentences read together, he wrote, "reflects Congress' clear and unambiguous intent to retain state law remedies except in the event of a conflict between those remedies and the CDA."[420]

Because Section 230 did not expressly preempt state law claims, the statute would block Zeran's claims only if Ellis found "field" or "conflict" preemption.[421] . Field preemption "occurs when federal law occupies a 'field' of regulation 'so comprehensively that it has left no room for supplementary state legislation.'"[422] Ellis concluded that, by passing Section 230, Congress had no intention to occupy the entire field of internet regulation, "but rather to eliminate obstacles to the private development of blocking and filtering technologies capable of restricting inappropriate online content."[423]

Conflict preemption occurs when "Congress enacts a law that imposes restrictions or confers rights on private actors; a state law confers rights or imposes restrictions that conflict with the federal law; and therefore the federal law takes precedence and the state law is preempted."[424] Ellis found that Section 230 preempted Zeran's claim against AOL because it conflicted with Section 230.[425] Because he concluded that distributor liability is a type of publisher liability, Ellis reasoned that "Zeran's attempt to impose distributor liability on AOL is, in effect, an attempt to have AOL treated as the publisher of the defamatory material. This treatment is contrary to §

---

416. *See* SYKES & VANATKO, *supra* note 396, at 2.

417. Zeran v. Am. Online, Inc., 958 F. Supp. 1124, 1130 (E.D. Va. 1997) (quoting 29 U.S.C. § 1144(c)(1)), *aff'd*, 129 F.3d 327 (4th Cir. 1997).

418. *Id.*

419. 47 U.S.C. § 230(e)(3).

420. *Zeran*, 958 F. Supp. at 1131.

421. *Id.*; Murphy v. NCAA, 138 S. Ct. 1461, 1479-80 (2018).

422. *Id.* at 1480 (quoting R.J. Reynolds Tobacco Co. v. Durham Cnty., 479 U.S. 130, 140 (1986)).

423. *Zeran*, 958 F. Supp. at 1131.

424. *Murphy*, 138 S. Ct. at 1480.

425. *Zeran*, 958 F. Supp. at 1132.

230(c)(1) of the CDA and, thus, Zeran's claim for negligent distribution of the notice is preempted."[426]

Alternatively, Ellis also concluded that Section 230 preempted Zeran's tort claim because it conflicted with Section 230's purposes "to encourage the development of technologies, procedures and techniques by which objectionable material could be blocked or deleted either by the interactive computer service provider itself or by the families and schools receiving information via the Internet."[427]

The Fourth Circuit's affirmance of Judge Ellis gave less attention to the doctrine of preemption, writing that "Congress' command is explicitly stated. Its exercise of its commerce power is clear and counteracts the caution counseled by the interpretive canon favoring retention of common law principles."[428] While the Fourth Circuit did not elaborate on this conclusion, it at least suggested that the court views Section 230(e)(3) as an express preemption provision.

Since then, courts have generally accepted that Section 230 preempts state common law and statutory claims, but they rarely delve deeply into preemption doctrine. In 2001, one of the first Section 230 cases after *Zeran*, the Florida Supreme Court, in ruling that Section 230 preempted a different negligence lawsuit against AOL, adopted the *Zeran* district court's reasoning that conflict preemption applied.[429] And in 2013, a federal judge in Tennessee concluded that Section 230 triggered both express preemption and conflict preemption.[430] While courts and commentators often disagree about whether Section 230 applies to particular types of claims, the disputes typically focus on whether the online platform materially contributed to the content at issue,[431] or whether the claim actually treats the platform as a publisher or speaker of third-party content.[432] There is no disagreement, however, about whether Section 230 can preempt state law.

In short, Section 230's history over the past quarter century instructs us that Congress has great leeway to preempt state defamation claims. Congress has the power, as granted in the Supremacy Clause and interstate Commerce Clause, to abrogate the ability of state courts to impose consequences for allegedly defamatory statements. To be sure, Congress's power is not absolute; it would impose these limits on defamation cases under the Commerce Clause. Theoretically, Congress might have trouble preempting a purely intrastate defamation claim. But to the extent that an allegedly libelous statement is circulated across state lines via the internet or any other medium,

426. *Id.* at 1133.
427. *Id.* at 1134.
428. *Zeran*, 129 F.3d at 334.
429. Doe v. Am. Online, Inc., 783 So. 2d 1010, 1016 (Fla. 2001).
430. Backpage.com, LLC v. Cooper, 939 F. Supp. 2d 805, 823-24 (M.D. Tenn. 2013).
431. *See, e.g.*, Fair Hous. Council of San Fernando Valley v. Roommates.com, LLC, 521 F.3d 1157, 1168 (9th Cir. 2008) (en banc).
432. *See, e.g.*, Lemmon v. Snap, Inc., 995 F.3d 1085, 1087 (9th Cir. 2021).

there is a very strong argument that any resulting dispute affects interstate commerce and is subject to congressional regulation.[433]

VI. THE FREEDOM OF SPEECH AND PRESS ACT

Now that we have reviewed the history of *Sullivan* and the Court's subsequent cases, as well as the rising tide of defamation lawsuits and how Congress might use its Commerce Clause powers to preempt libel law, we can propose an appropriate statutory fix to insulate the principles that *Sullivan* sought to protect. Alexander Meiklejohn's belief that speech and press protections are necessary for self-governance set the stage for *Sullivan*. The need for democracy-promoting speech safeguards has not dissipated in the past half-century. If anything, the rising tide of authoritarianism makes these protections more vital than ever. If the Supreme Court were to overturn *Sullivan*, it would be all the more difficult for the media and other speakers to investigate and criticize those in power.

Our proposal uses preemption to codify not only *Sullivan*'s protections, but the subsequent Supreme Court opinions that built on *Sullivan*. If the Supreme Court overturns *Sullivan*, the fate of these other precedents also is at stake, as they rely heavily on the 1964 opinion. The proposal, thus, recognizes that, were Congress to move to protect *Sullivan*, Congress should also take the opportunity to expand the protections that *Sullivan* and its progeny provide by incorporating other limitations that individual justices have advocated for in those cases—even if they ultimately did not obtain a majority for those positions. Indeed, our review of the rise of defamation cases demonstrates that many challenges face publishers despite *Sullivan*'s protections.

The full text of the proposal is in Appendix A. This Section summarizes the key provisions and points to parallels in the Court's First Amendment jurisprudence that inspired some of the proposal's provisions. The proposal, as explained in Section 1, is titled the "Freedom of Speech and Press Act."

Section 2 provides congressional findings that summarize the purpose of the statute. The section, based on the SPEECH Act of 2010 that passed by unanimous consent,[434] makes clear that the purpose is to codify the protections of *Sullivan* and its progeny. For instance, Section (2)(b) recognizes the nation's "profound national commitment to the principle that debate on public issues should be uninhibited, robust, and wide-open, and that it may well include vehement, caustic, and sometimes unpleasantly sharp attacks,"[435] language directly from the *Sullivan* opinion. Section 2 also explains how the

433. *See* United States v. MacEwan, 445 F.3d 237, 244 (3d Cir. 2006) ("[W]e conclude that because of the very interstate nature of the Internet, once a user submits a connection request to a website server or an image is transmitted from the website server back to user, the data has traveled in interstate commerce."); United States v. Carroll, 105 F.3d 740, 742 (1st Cir. 1997) ("Transmission of photographs by means of the Internet is tantamount to moving photographs across state lines and thus constitutes transportation in interstate commerce."); *see also* Gutierrez, *supra* note 384, at 44-47 (discussing preemption).

434. SPEECH Act of 2010, Pub. L. No. 111-223, 124 Stat. 2380.

435. N.Y. Times Co. v. Sullivan, 376 U.S. 254, 270 (1964).

threat of weaponized defamation lawsuits, particularly those brought by public officials and figures like those previously reviewed, can "inhibit other expression that might otherwise have been spoken, written, or published but for the fear of the lawsuit." The findings section is intended to leave no doubt among judges that the Freedom of Speech and Press Act is intended to codify *Sullivan* and its progeny and provide nationwide minimum protections for defamation defendants.

The Act also recognizes that some state jurisdictions might be less protective of speakers, which in the case of overruling *Sullivan*, might lead to drastically different rules state to state. Indeed, while Thomas has argued that the "States are perfectly capable of striking an acceptable balance between encouraging robust public discourse and providing a meaningful remedy for reputational harm," the sociopolitical history of *Sullivan* itself reveals precisely the opposite.[436] Alabama used libel law, unrestricted by the First Amendment, not to redress harm to reputation but to wage political warfare against unpopular speech and unpopular speakers. With other states, like Florida, presently moving to challenge *Sullivan* by laws drawing its approach into doubt, the recognition that some states are likely to use defamation law as a political cudgel is important.

Section 3 establishes the minimum level of fault that a plaintiff must establish before imposing liability for defamation. The bill's fault standard improves upon *Gertz*'s public figure/private figure distinction, which has long received criticism for its unpredictability.[437] Rather than forcing speakers to guess in advance whether a subject might be viewed as a public or private figure, the bill adopts the more predictable *Rosenbloom* plurality view on whether the underlying matter is of public concern.[438] (A similar focus was adopted by the Court in *Hepps* to determine when a plaintiff must bear the burden of proving falsity; despite criticism over the malleability of a public concern standard in *Rosenbloom*, courts have shown that they are perfectly capable of applying this standard.[439]) The bill broadly defines "public concern" as "any subject other than a purely private concern, including all matters of political, social, or other concern to the community," further alleviating any such difficulties in determining the contours of a matter of public concern.

Under the proposed statutory text, if the defamation lawsuit relates to a matter of public concern, the plaintiff must meet the actual malice standard of *Sullivan*, as interpreted in *St. Amant*: pleading and ultimately proving by clear

436. McKee v. Cosby, 139 S. Ct. 675, 682 (2019) (Thomas, J., concurring in denial of certiorari).

437. *See, e.g.*, Mark D. Walton, *The Public Figure Doctrine: A Reexamination of* Gertz v. Robert Welch, Inc. *in Light of Lower Federal Court Public Figure Formulations*, 16 N. ILL. U.L. REV. 141, 173 (1995) ("While the failure to delineate clear, rigid rules provides a needed degree of flexibility in many areas of the law, the inability of the media to accurately predict whether a statement will receive First Amendment protection prior to publication results in the suppression of information, a result that the *New York Times* Court sought to prevent.").

438. Rosenbloom v. Metromedia, Inc., 403 U.S. 29 (1971) (plurality opinion), *abrogated by* Gertz v. Robert Welch, Inc., 418 U.S. 323 (1974).

439. *Id.* at 79 (Marshall, J., dissenting).

and convincing evidence that the statement was made with knowledge of its falsity or that the defendant had a high degree of awareness of probable falsity.[440] If a lawsuit does not relate to a matter of public concern, at minimum a plaintiff must prove fault by a preponderance of evidence, a standard much like what *Gertz* required of private figure plaintiffs.[441] Codifying the two levels of fault based on public concern as opposed to the plaintiff's status as a public official or figure has support in recent precedent. In 2020, New York amended its anti-SLAPP law to require plaintiffs to demonstrate actual malice in defamation cases connected to "an issue of public interest."[442]

Section 4 ensures that states do not place the burden on defendants to prove the truth. *Sullivan*, followed by *Gertz* and *Hepps*, substantially changed earlier defamation law regimes by placing the burden of proving falsity on the plaintiffs.[443] The bill prevents state defamation laws from reverting to the pre-*Sullivan* standards placing the burden of proving truth on the defendant by requiring plaintiffs in lawsuits regarding matters of public concern to establish falsity by clear and convincing evidence. For other cases, the plaintiffs still have the burden of establishing falsity by a preponderance of the evidence.

The burden of proof is more than a legal technicality; in many defamation lawsuits, it could be dispositive. Consider a hypothetical defamation lawsuit that a city council member files against a citizen who posted on Facebook that she observed the council member taking cash from a local developer. If the defendant has the burden of proving that the statement was true, she might have a tough time establishing that the politician did, indeed, take the cash (unless she had a photograph, witnesses, bank statements, or other corroborating evidence). But if the plaintiff has the burden of proving falsity, the council member will face a heavy lift to establish that no cash changed hands.

Section 4 also incorporates the *Milkovich* standard and ensures that no state can impose liability for the expression of pure opinion, which the statute broadly defines as "any expression of opinion not subject to objective proof relating to matters of personal taste, aesthetics, criticism, religious beliefs, moral convictions, political views, or social theories." The provision only allows liability if the opinion alleges undisclosed defamatory facts as its basis, a standard that aligns with *Milkovich*.[444] This would prevent, for instance, the city council member suing the Facebook critic for posting that he is "the most awful person ever elected to city council." As defined in the statute, such a statement is a matter of pure opinion.

---

440. St. Amant v. Thompson, 390 U.S. 727, 731 (1968).

441. *Gertz*, 418 U.S. at 347.

442. S. 52A, 2019-2020 Leg., Reg. Sess. (N.Y. 2020) (enacted).

443. Floyd Abrams, *Preface* to MEDIA L. RES. CTR., NEW YORK TIMES V. SULLIVAN: THE CASE FOR PRESERVING AN ESSENTIAL PRECEDENT, at iv (2022), https://live-medialaw.pantheonsite.io/wp-content/uploads/2022/03/nytsullivanwhitepaper-1.pdf [https://perma.cc/47JV-LELG].

444. Milkovich v. Lorain J. Co., 497 U.S. 1, 20-21 (1990).

Section 5 sets limits on damages to avoid the chilling effect on speech that has concerned the Court since 1964.[445] The Section again adopts the *Hepps* focus on matters of public concern, while also recognizing the concerns about special and punitive damages that the Court recognized in *Gertz*.[446] The bill also reflects the concerns about the chilling effect of punitive and presumed damages that Marshall recognized in his *Rosenbloom* dissent: "The unlimited discretion exercised by juries in awarding punitive and presumed damages compounds the problem of self-censorship that necessarily results from the awarding of huge judgments."[447]

To address this issue, the bill would first require a showing by clear and convincing evidence of special damages in defamation suits arising from matters of public concern. Thus, plaintiffs must demonstrate provable pecuniary losses traceable to the alleged defamation. It would also cap punitive damages in those cases at three times the total compensatory damages. And, it would do away with presumed damages in such cases. This would make the threat of damages that animated the majority opinion in *Gertz* less likely to chill speech by limiting the quantum of them and making them also more difficult to prove.

For lawsuits arising from matters that are not of public concern, punitive and presumed damages would only be available with a showing, by clear and convincing evidence, that the statement was made with actual malice as defined by *Sullivan* and *St. Amant.* The bill aims to strike a balance by allowing defendants to recover damages—and even punitive damages in some cases—but capping those awards to ensure that they are tied more closely to the harms that the plaintiffs suffered and not merely to the desire of a judge or jury to punish the defendant.

Section 6 prevents the United States from returning to the days of seditious libel prosecutions by prohibiting criminal liability for defamatory statements. This is more than a theoretical concern; about half the states have laws on the books that allow for imprisonment, fines, or other criminal liability for defamatory statements.[448] Amid growing concern about the rising tides of authoritarianism in the United States, it is vital that Congress prevent a state legislature and governor from reinvigorating their criminal libel laws to punish dissenters. And while the Court in *Garrison* applied *Sullivan*'s actual malice rule as a limit on criminal libel law,[449] as our review of the recent weaponization of libel law has shown, that limitation provides little protection from a law enforcement investigation that is without, in the first instance, judicial intervention. We thus take the absolutist position advanced by Black and Douglas in *Garrison*: "[T]he First Amendment, made applicable to the States by the Fourteenth, protects every person from having a State or the Federal Government fine, imprison or assess damages against him when he

445. *Sullivan*, 376 U.S. at 300 (Goldberg, J., concurring).

446. *Gertz*, 418 U.S. at 347.

447. *Rosenbloom*, 403 U.S. at 84 (Marshall, J., dissenting).

448. *See Map of States with Criminal Laws Against Defamation*, AM. CIVIL LIBERTIES UNION, https://www.aclu.org/issues/free-speech/map-statescriminal-laws-against-defamation [https://perma.cc/DXS5-ASWW] (last visited Nov. 7, 2022).

449. Garrison v. Louisiana, 379 U.S. 64, 67 (1964).

has been guilty of no conduct, other than expressing an opinion, even though others may believe that his views are unwholesome, unpatriotic, stupid or dangerous"[450]

Section 7 establishes the preemptive effect of the Freedom of Speech and Press Act, adopting language directly from Section 230(e)(3), modified only slightly. The established—and unquestioned—caselaw regarding Section 230's preemption of state laws provides a solid basis for similar preemption of state defamation laws that do not meet the minimum requirements of this bill.[451] This preemption would create a floor for free speech protections, and states still would be free to provide even greater protections for defamation defendants. For instance, a state might choose to require plaintiffs to establish clear and convincing evidence of actual malice in all defamation claims, no matter if they involve matters of public concern. Likewise, a state could adopt the position that Black, Douglas, and Goldberg took in *Sullivan* and bar all defamation claims by public officials,[452] even if they established actual malice by clear and convincing evidence.

VII. CONCLUSION

In an ideal world, it would be unnecessary to codify and bolster a half-century of First Amendment precedent into a federal statute. But we are not in an ideal world. We are in a world in which the Supreme Court will radically change precedent in the name of originalism or textualism or pragmatism or whatever other theory suits its goals. We are in a world in which at least two Supreme Court Justices have called for their colleagues to reconsider *Sullivan*.[453] We are in a world in which politicians and powerful corporations weaponize libel laws to stifle criticism.

Overturning *Sullivan* would do more than eliminate the actual malice requirement for public official plaintiffs. It would undercut all First Amendment protections in defamation cases. It would open the door for state legislators and judges to enact oppressive punishments for those who had the gall to criticize the powerful, much like Alabama did in the 1960s. Without First Amendment protections, legislators and judges could give the subjects of criticism the ability to drive critics into bankruptcy, even with a terribly weak case. There would be no limits to the States' use of civil and criminal defamation laws as a tool to silence the opposition. Democracy would be worse for it.

But Congress could prevent such harms and provide journalists, social media posters, and all other speakers the assurances they need to speak freely. When the Supreme Court overturned *Roe v. Wade* in 2022, Justice Brett Kavanaugh wrote a concurrence that suggested legislators at either the state *or* federal level can determine what protections for abortion are available to the public. "The Constitution is neutral and leaves the issue for the people and

450. *Id.* at 79 (Black, J., concurring).
451. *See* KOSSEFF, *supra* note 401, at 3.
452. *See, e.g.*, *Sullivan*, 376 U.S. at 300 (Goldberg, J., concurring).
453. *See supra* Part III.

their elected representatives to resolve through the democratic process in the States or Congress—like the numerous other difficult questions of American social and economic policy that the Constitution does not address."[454] Just as Congress could codify *Roe*, it also could codify *Sullivan* and the other First Amendment defamation cases.

Although we believe that the First Amendment does, in fact, limit defamation liability, it is possible that five Justices will disagree. If that happens, a law such as the Freedom of Speech and Press Act would preserve the legal protections that have defined modern speech and journalism. And, it would protect freedom of speech and the press for future generations to come.

VIII. APPENDIX: TEXT OF THE FREEDOM OF SPEECH AND PRESS ACT

An Act

To provide national protections for freedom of speech and press.

Be it enacted by the Senate and House of Representatives of the United States of America in Congress assembled,

Section 1. SHORT TITLE

This Act may be cited as the "Freedom of Speech and Press Act."

Section 2. FINDINGS

Congress finds the following:

(a) The freedom of speech and of the press is enshrined in the First Amendment to the Constitution and is necessary to promote the vigorous dialogue necessary to shape public policy in a representative democracy.

(b) Our nation has a profound national commitment to the principle that debate on matters of public concern should be uninhibited, robust, and wide-open, and that it may well include vehement, caustic, and sometimes unpleasantly sharp attacks.

(c) Some persons are obstructing the free speech and free press rights of United States citizens and frustrating this commitment through the weaponization of defamation by seeking out state jurisdictions that do not provide the full extent of free-speech, free-press protections owed to United States citizens, and suing in those jurisdictions.

(d) These retaliatory lawsuits not only suppress the free speech and press rights of the defendants in the lawsuit, but inhibit other expression that

454. Dobbs v. Jackson Women's Health Org., 142 S. Ct. 2228, 2305 (2022) (Kavanaugh, J., concurring).

might otherwise have been spoken, written, or published but for the fear of a lawsuit.

(e) The internet and the mass distribution of media interstate, including through, among other channels, broadcast, cable, and satellite services, also create the danger that one State's unduly restrictive defamation law will affect freedom of speech and press worldwide on matters of public concern.

(f) This country's debate on matters of public concern will be fostered by adopting national standards and requirements relating to the law of defamation, as defined herein.

Section 3. FAULT REQUIRED

(a) In any defamation lawsuit relating to a matter of public concern, no State shall impose liability absent a plaintiff pleading and ultimately proving by clear and convincing evidence that the statement was made with knowledge of its falsity or that the defendant had a high degree of awareness of probable falsity.

(b) In any defamation lawsuit that does not relate to a matter of public concern, no State shall impose liability without a plaintiff pleading and ultimately proving by a preponderance of evidence a defendant's fault.

Section 4. FALSITY REQUIRED

(a) In any defamation lawsuit relating to a matter of public concern, no State shall impose liability absent a plaintiff pleading and ultimately proving by clear and convincing evidence that the challenged statement is materially false.

(b) In any defamation lawsuit not relating to a matter of public concern, no state shall impose liability absent a plaintiff pleading and ultimately proving by a preponderance of the evidence that the challenged statement is materially false.

(c) An opinion can be actionable only if it implies undisclosed defamatory facts as a basis of the opinion or, alternatively, is based on disclosed but false facts. In any defamation lawsuit, no state shall impose liability for a pure opinion nor for an opinion based on disclosed, substantially true facts.

Section 5. LIMITATIONS ON DAMAGES

(a) In any defamation lawsuit relating to a matter of public concern, no State shall impose liability absent a plaintiff pleading and ultimately proving by clear and convincing evidence special damages caused by the allegedly defamatory statement.

(b) In any defamation lawsuit relating to a matter of public concern, no State shall provide for an award of punitive damages that exceeds three times the total compensatory damages awarded.

(c) In any defamation lawsuit relating to a matter of public concern, no State shall provide for an award of presumed damages.

(d) In any defamation lawsuit not relating to a matter of public concern, no State shall provide for an award of punitive damages absent a plaintiff pleading and ultimately proving by clear and convincing evidence that the statement was made with knowledge of its falsity or that the defendant had a high degree of awareness of probable falsity.

(e) In any defamation lawsuit not relating to a matter of public concern, no State shall provide for an award of presumed damages absent a plaintiff pleading and ultimately proving by clear and convincing evidence that the statement was made with knowledge of its falsity or that the defendant had a high degree of awareness of probable falsity.

Section 6. PROHIBITION ON CRIMINAL LIBEL

No State shall impose criminal liability based solely on the dissemination of an allegedly false and defamatory statement or statements.

Section 7. EFFECT ON OTHER LAWS

(a) NO EFFECT ON INTELLECTUAL PROPERTY LAW – Nothing in this section shall be construed to limit or expand any law pertaining to intellectual property.

(b) STATE LAW – Nothing in this law shall be construed to prevent any State from enforcing any State law that is consistent with or provides protections for freedom of speech and of the press in excess of those provided by this law. No cause of action may be brought and no liability may be imposed under any State or local law that is inconsistent with this law.

Section 8. DEFINITIONS

(a) The term "defamation" means any action or other proceeding for defamation, libel, slander, or similar claim alleging that forms of speech are false, have caused damage to reputation or emotional distress, have presented any person in a false light, or have resulted in criticism, dishonor, or condemnation of any person.

(b) "Materially false" shall mean a statement that would have a different effect on the mind of the reader, listener, viewer, or other recipient from that which the pleaded truth would have produced.

(c) "Public concern" shall be construed broadly, and shall mean any subject other than a purely private concern, including all matters of political, social, or other concern to the community.

(d) "Pure opinion" shall be construed broadly, and shall mean any expression of opinion not subject to objective proof relating to matters of personal taste, aesthetics, criticism, religious beliefs, moral convictions, political views, or social theories. Pure opinion includes resort to rhetorical hyperbole, satire, parody, and other forms of criticism that a reasonable reader would understand as not intending to convey actual facts.

Famously Fake: Using the Law to Reverse the Demise of Social Media Credibility

Delaney Dunn*

TABLE OF CONTENTS

* J.D., May 2023, The George Washington University Law School. Thank you to Professor Meredith Rose for her assistance in drafting this Note, to the other professors and students who are too many to name that helped along the way, and to the professionals I met in graduate school who let me pick their brains about the practical difficulties they faced in this industry.

I. INTRODUCTION

Caroline Calloway was arguably the first social media influencer.[1] She amassed hundreds of thousands of followers in the early days of Instagram by sharing her personal experiences, thoughts, and feelings about her life as an American traveling around Europe and attending the University of Cambridge.[2] This was a different era of social media, when sharing such intimate details with the public was shocking and the concept of social media influencers was so new that Calloway's classmates and acquaintances treated what Calloway was doing like it was a joke.[3] By spring 2015, Calloway had accumulated 300,000 followers on Instagram.[4] While this would be a relatively small following today, in 2015, it made her one of the most influential people on the Internet. A year later, Calloway graduated Cambridge with Instagram fame in one hand and a $500,000 book deal in the other.[5] She set the standard and blazed a trail for all future social media influencers to follow. The only problem was it was all—or at least mostly—fake.[6]

Calloway did not initially want to be an influencer.[7] Calloway always wanted to be a writer, and when she was rejected by publishers because no one wanted to publish or read the memoir of a nobody (even a rich nobody), she decided to become a somebody.[8] Fortunately for Calloway, there was a new social media app on the market called Instagram.[9] She recognized the potential Instagram had to offer and decided to stamp her name on it.[10] With the help of her best friend Natalie Beach, Calloway posted high-quality pictures and detailed captions about her life.[11] Her follower count soared, and thanks to Calloway, it became clear to the world that there was a market for non-traditional celebrities to build influence and fame on social media.[12] Eventually, Calloway was offered the book deal of her dreams, but she became so caught up in her Instagram fame that she struggled to finish the book that publishers had offered her half a million dollars to write.[13] Beach,

---

1. *See* Harling Ross, *Was Caroline Calloway the First Instagram Influencer?*, REPELLER (June 20, 2018), https://repeller.com/caroline-calloway-interview/ [https://perma.cc/ZSY6-NB9F].
2. *Id.*; Jacob Shamsian et al., *How Caroline Calloway Went from Instagram Influencer with a $500,000 Book Deal to the Creator of Her Personal 'Fyre Festival'*, INSIDER (Sept. 11, 2019, 11:10 AM), https://www.insider.com/caroline-calloway-book-deal-instagram-career-2019-1 [https://perma.cc/SY2A-CCW4].
3. Shamsian et al., *supra* note 2.
4. *Id.*
5. *Id.*
6. *See generally id.*; s*ee generally* Natalie Beach, *I Was Caroline Calloway*, CUT (Sept. 10, 2019), https://www.thecut.com/2019/09/the-story-of-caroline-calloway-and-her-ghostwriter-natalie.html [https://perma.cc/YL96-YVKD].
7. Beach, *supra* note 6.
8. *Id.*
9. *Id.*
10. *Id.*
11. *Id.*
12. *Id.*
13. Beach, *supra* note 6.

always the helpful best friend, started writing the book for Calloway.[14] It was not long after this that Calloway finally confessed something to Beach—Calloway's fame had not grown organically from years of posting interesting, aspirational content. [15] It grew because Calloway had been buying tens of thousands of followers to boost her numbers in the hopes of convincing publishers that people cared what she had to say.[16] Calloway bought her way to being one of the first truly famous Internet stars and to a lucrative book deal.[17] Not long after Beach found out, publishers rescinded Calloway's book deal without any public explanation, and Calloway's influence on social media began to wane.[18] It was only years later, when Beach decided to come out and tell her side of the Caroline Calloway story, that people found out why.[19] Today, Calloway is famous as a woman who took advantage of her fans and the companies she promoted, and lost everything because of it.[20]

She may have been the first, but Caroline Calloway was by no means the last influencer to try to buy their way to fame and fortune.[21] Today, it is even easier for influencers like Calloway to abuse the system.[22] The robust advertising agencies and marketing departments of old are being replaced with the whims of a single individual, often a teenager, who does not work for or have any longstanding relationship with the brand.[23] Companies have made this change out of necessity.[24] The repercussions of buying fake

14. *Id.*
15. *Id.*
16. *Id.*
17. *Id.*; Lauren Frias, *Instagram Influencer Caroline Calloway Bought Followers and Created Her Own Fan Base to Sell Her Unwritten Memoir, Her Ghostwriter Claims in an Explosive New Essay*, INSIDER (Sept. 10, 2019, 10:18 PM), https://www.insider.com/instagram-influencer-caroline-calloway-bought-followers-created-own-fan-base-2019-9 [https://perma.cc/KYX2-EHYJ].
18. Frias, *supra* note 17.
19. Beach, *supra* note 6.
20. Shamsian et al., *supra* note 2.
21. Frias, *supra* note 17; *see* Gil Appel et al., *The Future of Social Media Marketing*, J. ACAD. MKTG. SCI. 79, 89 (2019); *see* Tom Huddleston Jr., *How Instagram Influencers Can Fake Their Way to Online Fame*, CNBC (Feb. 4, 2021, 2:58 PM), https://www.cnbc.com/2021/02/02/hbo-fake-famous-how-instagram-influencers-.html [https://perma.cc/QGL9-BBSU].
22. *See* Appel et al., *supra* note 21; s*ee Demand for Fake Instagram Followers Shot Up 71% This Year*, MEDIAKIX, https://web.archive.org/web/20210418003958/https://mediakix.com/blog/fake-instagram-followers-growing-demand/ [https://perma.cc/37A7-DPSL] (last visited Jan. 23, 2022) [hereinafter *Demand for Fake Instagram Followers*].
23. *See* Shareen Pathak, *Brands Are Using Influencers like Ad Agencies*, DIGIDAY (May 24, 2017), https://digiday.com/marketing/brands-using-influencers-like-ad-agencies/ [https://perma.cc/7KUQ-XQ7M].
24. *See* Danielle Wiley, *How to Use Influencers as Your Brand's Secret Weapon for the Next Normal*, FORBES (Sept. 28, 2021, 7:00 AM), https://www.forbes.com/sites/forbesagencycouncil/2021/09/28/how-to-use-influencers-as-your-brands-secret-weapon-for-the-next-normal/ [https://perma.cc/L3DY-P9SD].

followers are usually limited to diminished credibility, though people like Caroline Calloway prove that the payoffs can be massive.[25]

Demand for fake followers is growing every day, and social media companies have done comparatively little to curb the rampant use of fake followers on their platforms.[26] The truth of the matter is social media companies do not *want* to get rid of fake accounts.[27] One study conducted by Ars Technica found that Facebook leaves up 95% of fake accounts on their social media platforms, even after those accounts are reported.[28] Twitter has even been known to verify fake accounts as celebrities or influencers to such an extent that they recently had to suspend their verification process entirely.[29] On the rare occasion social media companies have made attempts to curb fake followers or engagement, they have always taken action against websites selling fake followers rather than influencers purchasing them.[30] Facebook Inc., now called Meta, has personally filed several lawsuits against companies selling fake Instagram followers, which would be admirable if any of them had amounted to anything.[31] Only one company, who was previously the subject of a New York Times investigation, experienced any substantial

25. Frias, *supra* note 17; *See* Caroline Forsey, *Why You Shouldn't Buy Instagram Followers*, HUBSPOT (May 6, 2022), https://blog.hubspot.com/marketing/buy-instagram-followers [https://perma.cc/3E25-SZHS].

26. *See Demand for Fake Instagram Followers*, *supra* note 22; Kate Cox, *Social Media Platforms Leave 95% of Reported Fake Accounts Up, Study Finds*, ARS TECHNICA (Dec. 6, 2019, 2:42 PM), https://arstechnica.com/tech-policy/2019/12/social-media-platforms-leave-95-of-reported-fake-accounts-up-study-finds/ [https://perma.cc/TFA3-SMVY].

27. *See* Jack Morse, *Why Social Media Companies Won't Kill Off Bots*, MASHABLE (Feb. 6, 2018), https://mashable.com/article/facebook-instagram-twitter-bots [https://perma.cc/JCR9-KZUB].

28. *See* Cox, *supra* note 26.

29. *See* Sophie Webster, *Twitter to Halt Its Verification Process After Several Fake Accounts Got Verified*, TECH TIMES (Aug. 13, 2021, 6:08 PM), https://www.techtimes.com/articles/264121/20210813/twitter-halt-verification-process-several-fake-accounts-verified.html [https://perma.cc/6YUG-THMU].

30. *Facebook Sues over Sales of Fake Accounts, Likes and Followers*, CBS NEWS (Mar. 1, 2019, 9:16 PM), https://www.cbsnews.com/news/facebook-lawsuit-over-sales-of-fake-accounts-likes-and-followers-china/ [https://perma.cc/M42V-Y49Z].

31. *See* Paul Grewal, *Cracking Down on the Sale of Fake Accounts, Likes and Followers*, META (Mar. 1, 2019), https://about.fb.com/news/2019/03/sale-of-fake-accounts-likes-and-followers/ [https://perma.cc/Q4YL-PWG2]; Jessica Romero, *Taking Action Against Fake Engagement and Ad Scams*, META (Oct. 20, 2020), https://about.fb.com/news/2020/10/taking-action-against-fake-engagement-and-ad-scams/ [https://perma.cc/HAW2-Z248].

consequences for their actions.[32] The odds that social media companies will take action, and the odds that action will be effective, are negligible and insufficient to deter these websites.[33]

To prevent further damage, action must be taken to curb the demand. Brands experience immense damage, estimated at over $1 billion in 2019 alone, with that number growing as influencer marketing grows.[34] Given that social media platforms are not helping, brands themselves deserve to be empowered to mitigate the damage.[35] If we hope to see any real, concrete change, brands need to be given the opportunity to pursue the influencers who are creating the demand for fake followers. This can be done under current law by interpreting the definition of fraud to encompass the actions of these influencers.

This Note will begin with an explanation of how influencer marketing came to prominence in the marketing industry. It will then examine the role of influencers, why they have influence, and why marketers use them as a resource. The Note will then look at why and how influencers deceive marketers using bots and social pods and what damage it may do to brands. Following the Background section will be an analysis of how these influencers may be held liable. This section will consider whether influencers may be charged with fraud at the state level in order to curb their online falsifications. The elements of fraud—misrepresentation, knowledge of falsity, intent, and justifiable reliance—will each be considered, as will potential defenses influencers may raise to each, and factors that may limit which brands may utilize this method. The Analysis will also briefly discuss the resulting damages that are required for brands to experience for influencer activity to be considered fraud.

---

32. Nicholas Confessore et al., *The Follower Factory*, N.Y. TIMES (Jan. 27, 2018), https://www.nytimes.com/interactive/2018/01/27/technology/social-media-bots.html [https://perma.cc/C6ZT-TTMM]; *see* Press Release, FTC, Devumi Owner and CEO Settle FTC Charges They Sold Fake Indicators of Social Media Influencer; Cosmetics Firm Sunday Riley, CEO Settle FTC Charges That Employees Posted Fake Online Reviews at CEO's Direction (Oct. 21, 2019), https://www.ftc.gov/news-events/press-releases/2019/10/devumi-owner-ceo-settle-ftc-charges-they-sold-fake-indicators [https://perma.cc/GH75-6EW5] [hereinafter FTC Press Release] (levying $2.5 million settlement against Devumi CEO); Press Release, N.Y.S. Off. Att'y Gen., Attorney General James Announces Groundbreaking Settlement with Sellers of Fake Followers and 'Likes' on Social Media (Jan. 30, 2019), https://ag.ny.gov/press-release/2019/attorney-general-james-announces-groundbreaking-settlement-sellers-fake-followers [https://perma.cc/4NTD-BBF8] (announcing settlement that prohibits engaging in similar activity).

33. *See* FTC Press Release, *supra* note 32 (showing the FTC judgments and lawsuits in 2019); *see* Cox, *supra* note 26; Nicholas Confessore & Gabriel J.X. Dance, *On Social Media, Lax Enforcement Lets Impostor Accounts Thrive*, N.Y. TIMES (Feb. 20, 2018) https://www.nytimes.com/2018/02/20/technology/social-media-impostor-accounts.html [https://perma.cc/TJ95-2KXS].

34. Emma Grey Ellis, *Fighting Instagram's $1.3 Billion Problem – Fake Followers*, WIRED (Sept. 10, 2019, 8:00 AM), https://www.wired.com/story/instagram-fake-followers/ [https://perma.cc/8JFG-FSQ4] (showing the billions in damages brands have lost in advertising to fake followers).

35. *Id.*

II. BACKGROUND

A. Marketing Shifts

Since Caroline Calloway's rise to fame, there has been a significant shift in the marketing industry towards influencers.[36] The advent of digital marketing methods has been slowly pushing traditional marketing out of frame for years.[37] Digital marketing, or marketing using the Internet, has significant advantages over traditional methods such as newspapers and television. Digital marketers can easily track relevant data, such as how many people see an ad and with which ads consumers interact.[38] This data can help marketers customize ads to individual consumers and improve the overall quality of ads.[39]

Unfortunately, digital marketing has exacerbated some of the worst issues facing marketers. Over the last few decades, the number of ads the average consumer sees per day has risen to between 4,000 and 10,000; but as the number of ads has grown, the effectiveness of each has lessened.[40] Consumers are understandably burnt out.[41] They are tired of boring and irrelevant ads, have shorter attention spans, and treat marketers with a healthy amount of mistrust.[42] Most ads do not factor into the consumer decision-

36. *See* Yusha Charles, *Why Is Influencer Marketing So Important in 2021*?, LINKEDIN (Apr. 25, 2021), https://www.linkedin.com/pulse/why-influencer-marketing-so-important-2021-yusha-charles/ [https://perma.cc/H7JP-M3ME] (including a list of reasons brands have turned to influencers for help in the digital age).

37. *See generally* Hamza Shaban, *Digital Advertising to Surpass Print and TV for the First Time, Report Says*, WASH. POST (Feb. 20, 2019, 9:53 AM), https://www.washingtonpost.com/technology/2019/02/20/digital-advertising-surpass-print-tv-first-time-report-says/ [https://perma.cc/4NUZ-CQQK].

38. *See id.*

39. *Id.*

40. *See* Tara Drosset, *How Many Times Must You See an Ad to Actually Remember It?*, RED CROW MKTG. (July 7, 2021), https://www.redcrowmarketing.com/2021/07/07/how-many-times-must-you-see-an-ad-to-remember-it/ [https://perma.cc/GAR2-5MQ8] (discussing how people barely remember seeing ads, and now they require repetition to even register); s*ee* Sam Carr, *How Many Ads Do We See a Day in 2021?*, PPC PROTECT (Feb. 15, 2021), https://ppcprotect.com/blog/strategy/how-many-ads-do-we-see-a-day/ [https://perma.cc/W3Y8-N99E].

41. *See* Nishat Mehta, *Give Consumers the Ads They Want*, FORBES: COMM. COUNCIL (Mar. 7, 2018), https://www.forbes.com/sites/forbescommunicationscouncil/2018/03/07/give-consumers-the-ads-they-want/ [https://perma.cc/UNP8-7T62]; *see* Bill Lee, *Marketing Is Dead*, HARV. BUS. REV. (Aug. 9, 2012), https://hbr.org/2012/08/marketing-is-dead [https://perma.cc/2SXA-FYP6].

42. *See* Mehta, *supra* note 41; s*ee* Lee, *supra* note 41; *see* Kai Ryssdal, *Goldfish Have Longer Attention Spans Than Americans, and the Publishing Industry Knows It*, MARKETPLACE (Feb. 11, 2014), https://www.marketplace.org/2014/02/11/goldfish-have-longer-attention-spans-americans-and-publishing-industry-knows-it/ [https://perma.cc/J9JQ-PY4H] (commenting on consumer attention span and how it has forced even the book industry to adapt, shorten time spans between books, and shorten books themselves to get Americans to pick them up; also noting shortening of attention span from twelve to eight seconds in thirteen years).

making process, and marketers are essentially paying for expensive background noise.[43] This is where influencers can be of great benefit.

B. Influencers

Influencers are social media users who have the ability to influence the decision-making processes of their audience.[44] Influencers can exist at a macro or micro scale; users with as few as a thousand followers are referred to as "micro-influencers," while household names with hundreds of millions of followers are considered "macro-influencers."[45] Most celebrities have large fanbases who listen to what they have to say, and while many of them could likely be considered influencers based on their persuasive power, the name usually only applies to those who became popular first and foremost for their activity on social media.[46]

1. Growing Influence

Influencers, regardless of the size of their audience, have to spend time investing in a relationship with their followers.[47] Influencers build a personal brand by making followers feel like friends, which leads their followers to trust and value the influencer's opinions.[48] This is one manifestation of "parasociality," a psychological concept referring to one-sided relationships where one party falsely perceives or misinterprets the existence of a friendship or relationship with the other.[49] This most commonly exists between celebrities and fans who believe they have a personal relationship

43. Debora Bettiga & Lucio Lamberti, *Future-Oriented Happiness: Its Nature and Role in Consumer Decision-Making for New Products*, FRONTIERS PSYCH., May 15, 2020, at 1, 5-6 (referring to ads as background noise due to lack of emotional response to them by consumers).

44. Werner Geyser, *What Is an Influencer? – Social Media Influencers Defined [Updated 2022]*, INFLUENCER MKTG. HUB (July 27, 2022), https://influencermarketinghub.com/what-is-an-influencer/ [https://perma.cc/FW6K-Y9D5].

45. *Id.*

46. Paul Jankowski, *Not All Influencers Are Celebrities . . . Not All Celebrities Are Influencers, Part 2*, FORBES (Mar. 5, 2021, 10:52 AM), https://www.forbes.com/sites/pauljankowski/2021/03/05/not-all-influencers-are-celebritiesnot-all-celebrities-are-influencers-part-2/ [https://perma.cc/7FZH-6FAG].

47. Steven Woods, #Sponsored: The Emergence of Influencer Marketing (2016) (B.S. thesis, University of Tennessee, Knoxville) (available at https://trace.tennessee.edu/cgi/viewcontent.cgi?article=3010&context=utk_chanhonoproj) [https://perma.cc/5BWE-FBB4].

48. Kate Ng, *Celebrity Endorsements Only Influence 4% of Shoppers, Survey Says*, INDEP. (Jan. 24, 2022, 9:46 AM), https://www.independent.co.uk/life-style/fashion/celebrity-influencer-endorsements-fashion-consumers-b1999000.html [https://perma.cc/74T3-7YWS] (discussing that micro-influencers having significantly more sway over their followers than celebrities or even mega-influencers because of the relationship they have with their followers); Lotte Bugge, *Why Influence Not Advertising is the Future for Brands*, TALKING INFLUENCE (Dec. 21, 2021) https://talkinginfluence.com/2021/12/21/why-influence-not-advertising-is-the-future-for-brands/ [https://perma.cc/2M6X-FBGQ].

49. *See generally* Chen Lou, *Social Media Influencers and Followers: Theorization of a Trans-Parasocial Relation and Explication of Its Implications for Influencer Advertising*, J. ADVERT., Mar. 2021, at 4.

with the celebrity, who they believe cares about them individually.[50] Parasocial dynamics are marked by significant power imbalances between the object of parasociality (the celebrity) and the perceiver of parasociality (the fan), and this is particularly prevalent on social media.[51] Influencers actively play into parasocial relationships to gain influence over their followers.[52]

As a result, even though followers *know* influencers are monetizing their presence, they also trust the relationship influencers have cultivated and generally believe influencers are less likely than companies or distant celebrities to lead them astray for their own personal gain.[53] In some cases, an influencer disclosing that they are making money off of their followers actually benefits them because it adds to the followers' perception that they are in an open, honest relationship.[54] Some influencers create content that emphasizes that they would never do wrong by their followers.[55] Influencers love to remind followers they do not partner with everyone that offers them money or exposure and that they only accept partnerships they genuinely believe would benefit their followers.[56] This adds to the perception that influencers are honest, trustworthy, and authentic in a way that traditional marketing has not been able to generate in recent years.[57]

2. Influencer Marketing

Marketers are trying to find new, creative ways to get their ads in front of consumers and to get those consumers to pay attention to their ads; to

---

50. *Id.*

51. *Id.*

52. *Id.*

53. *See* Joel Mathew, *Understanding Influencer Marketing and Why It Is So Effective*, FORBES: YOUNG ENTREPRENEUR COUNCIL (July 30, 2018, 8:00 AM), https://www.forbes.com/sites/theyec/2018/07/30/understanding-influencer-marketing-and-why-it-is-so-effective/ [https://perma.cc/Z56W-JZU2] (discussing how followers are not skeptical of influencer ads in the same way that they are skeptical commercials; followers are also less skeptical of influencers than celebrities who became famous); *see generally* Woods, *supra* note 47.

54. *See* Lou, *supra* note 49.

55. *See* Safiya Nygaard, *Trying Products That Asked to Sponsor Me (Not Sponsored)*, YOUTUBE (Oct. 7, 2018), https://www.youtube.com/watch?v=cY4e0uvp7uI [https://perma.cc/DRF9-ZA87] (trying brands that requested sponsorship that she refused or did not respond to because she only wants to promote products that she feels she knows enough about to recommend).

56. *See id.*

57. Lauren C., Comment to *Trying Products That Asked to Sponsor Me (Not Sponsored)*, YOUTUBE, https://www.youtube.com/watch?v=cY4e0uvp7uI&lc=UgwpovbiduZJVvBDdr14AaABAg [https://perma.cc/MDG7-ST2B] (where a comment with thousands of likes is praising Safiya for how responsible she is in what brands she chooses to do sponsorships with); Shraddha Kulshrestha, Comment to *Trying Products That Asked to Sponsor Me (Not Sponsored)*, YOUTUBE, https://www.youtube.com/watch?v=cY4e0uvp7uI&lc=UgxzpLr0MZxl8DRsFO94AaABAg [https://perma.cc/24Y7-DGGJ] (where a comment with thousands of likes is praising Safiya for being a "non-sponsored queen" on a video where she mentions multiple times that some of her other content is sponsored, just that she refused the sponsorships from the particular companies highlighted in this video).

accomplish this, marketers are turning to influencers for help.[58] Paying influencers to post ads is one of the most effective ways marketers achieve this, and it has several benefits over traditional marketing.[59] Influencers can leverage the trust they have built up with consumers to the brand's benefit.[60] Moreover, if a marketer does their research and chooses an influencer who projects values and interests similar to that of their brand, they will likely find that the ad connects more frequently with the influencer's followers than it would with the population at large who would see a traditional ad.[61] The growing reliance on influencers is evident in the fact that 86% of marketers plan to maintain or increase their influencer marketing budget in 2022.[62]

While marketers consider many factors when choosing which influencers to partner with and how much they are worth, the primary factors are follower count and engagement on posts.[63] The follower count of an influencer can give marketers an idea of the size of the audience an influencer has listening to them.[64] In traditional marketing, marketers used to have to estimate how many people would see an ad they posted in the paper or ran on television; on social media, marketers can look to follower counts for an exact or nearly exact number of people who will see whatever they pay the influencer to post.[65] Engagement, the other primary factor marketers take into account, is a measure of what percentage of followers like or comment on a post.[66] Engagement can give marketers an idea of how effective advertising with a particular influencer will be and thereby how much that influencer is worth.[67] Influencers with high engagement are believed to have more active followers who pay closer attention to their posts; this suggests that their followers care more about what the influencer has to say and will be more willing to take the influencer's opinion into account when they post an ad endorsing or vouching for a brand.[68]

---

58. *See* Carr, *supra* note 40.

59. *See generally* Mathew, *supra* note 53.

60. *See id.*

61. *See* Shannon Burton, *The Right Fit: How to Find Influencers for Your Brand's Marketing Campaign*, SPROUT SOC. (Jan. 25, 2021), https://sproutsocial.com/insights/how-to-find-the-right-influencers/ [https://perma.cc/MS4M-9ZDH].

62. Kristen Baker, *What Will Influencer Marketing Look Like in 2022?*, HUBSPOT (June 15, 2022), https://blog.hubspot.com/marketing/how-to-work-with-influencers [https://perma.cc/LSP6-Z7NL].

63. *See* Mathew, *supra* note 53 (factoring followers and reach into price of influencers).

64. *See* Geyser, *supra* note 44.

65. *Digital Versus Traditional Marketing: What Today's C-Suite Needs to Know*, WHARTON ONLINE (July 17, 2019), https://online.wharton.upenn.edu/blog/digital-versus-traditional-marketing/ [https://perma.cc/ES2Z-XKTX].

66. *See* Xabier Vicuña, *Choosing the Right Influencers: The Metrics That Matter*, FORBES: BUS. COUNCIL (Dec. 9, 2020, 7:20 AM), https://www.forbes.com/sites/forbesbusinesscouncil/2020/12/09/choosing-the-right-influencers-the-metrics-that-matter/?sh=5096df54709a [https://perma.cc/3NH6-KLLT] (explaining why engagement is as or more important than follower count).

67. *See id.*

68. *See id.*

3. Influencer Revenue

Influencers are now vying for billions of dollars in advertising revenue. The influencer marketing industry's estimated revenue for 2021 was approximately $13.8 billion, and that number is projected to surpass $15 billion in 2022.[69] Followers and engagement are critical when brands determine how much of that money an individual influencer is going to get.[70] For example, Charli D'Amelio, one of the five most followed influencers on the social media video app TikTok, charges between $100,000 and $250,000 per sponsored video she posts to her TikTok account.[71] There is no definitive formula for calculating how much a brand will pay a star, but the correlation between followers, engagement, and money is well documented.[72] Looking to Instagram, influencers with a million or more followers will typically make $7,500 or more per post, while those who fall between half a million to a million usually make around $5,000 per post, with value per post increasing or decreasing as follower and engagement counts increase or decrease.[73]

C. Deceiving Marketers

To increase their popularity (and their paychecks), many influencers will pay or trade for fake followers and fake engagement.[74] This practice is not limited to aspiring influencers but is common even among established influencers.[75] There are two ways to go about generating fake influence on social media: bot accounts and pods.

1. Bot Accounts

The most obvious method of falsifying activity on social media is bot accounts. Bot accounts are social media accounts that are not created or

69. Jacinda Santora, *Key Influencer Marketing Statistics You Need to Know for 2022*, INFLUENCER MKTG. HUB (Aug. 3, 2022), https://influencermarketinghub.com/influencer-marketing-statistics/ [https://perma.cc/JJX4-DYHU].

70. See Vicuña, *supra* note 66.

71. *See* Abram Brown & Abigail Freeman, *Top Earning TikTok-ers 2022: Charli and Dixie D'Amelio and Addison Rae Expand Fame — and Paydays*, FORBES (Jan. 7, 2022, 6:30 AM) https://www.forbes.com/sites/abrambrown/2022/01/07/top-earning-tiktokers-charli-dixie-damelio-addison-rae-bella-poarch-josh-richards/ [https://perma.cc/7YAX-LS5W] (showing the estimated earnings of the top stars on TikTok in 2021).

72. See Vicuña, *supra* note 66.

73. BRITTANY HENNESSY, INFLUENCER: BUILDING YOUR PERSONAL BRAND IN THE AGE OF SOCIAL MEDIA 141 (2018).

74. *See* Confessore et al., *supra* note 32.

75. *See* Abhinav Anand et al., *Influencer Marketing with Fake Followers* 2 (Indian Inst. of Mgmt. Bangalore, Working Paper No. 580, 2019), https://www.iimb.ac.in/sites/default/files/2019-02/WP%20No.%20580%20%28Revised%20Feb%202019%29.pdf [https://perma.cc/TG52-U8EM].

operated by real people.[76] They are created *en masse*, often through hacking, and are used for the purpose of following and engaging with content when their creator has paid for them to do so.[77]

There are websites devoted to selling bot accounts and bot engagement for every conceivable social media platform from Goodreads to Instagram.[78] Buying followers is an incredibly easy process. Websites only need to be provided with credit card information, a website, and a username, and they can make you famous overnight.[79] At a premium, websites selling this type of fake influence will go to great lengths to mask purchased followers and engagement.[80] Websites will make fake followers that look like real people and automated comments that, at a basic level, still appear related to the post they are left on.[81] Some websites will even trickle in purchased followers, creating the fake accounts and following the influencer slowly over time to mimic real growth.[82]

2. Pods

The other, less common way influencers gain unearned influence is through pods.[83] Pods vary greatly in size and rules, but the overall premise is much the same. Pods are private groups of real people where influencers trade likes and follows on other influencer's accounts for likes and follows in return.[84] Some pods are very narrow in scope, requiring influencers to all be from similar industries or share some common aesthetic, while others are essentially free-for-alls, letting in anyone who is willing to like and follow in

76. *See* Stefano Cresci et al., *Fame for Sale: Efficient Detection of Fake Twitter Followers*, DECISION SUPPORT SYS., Dec. 2015, at 56 ("*Fake followers* are those Twitter accounts specifically created to inflate the number of followers of a target account.").

77. *See* Masarah Paquet-Clouston et al., *Can We Trust Social Media Data? Social Network Manipulation by an IoT Botnet, in* #SMSOCIETY'17: PROC. FROM THE 8TH INT'L CONF. OF SOC. MEDIA & SOC'Y 1, 2 (2017).

78. *See, e.g.*, *Buy Goodreads Ratings, Reviews, Votes, Followers and Friends*, BADDHI SHOP, https://baddhi.shop/product/buy-goodreads-ratings/ [https://perma.cc/VYU4-QADY] (last visited Apr. 10, 2022); *6 Best Sites to Buy Instagram Likes Reviewed (2022)*, AMNY, https://www.amny.com/sponsored/buy-instagram-likes/ [https://perma.cc/GW9N-QTA4] (last visited Apr. 10, 2022).

79. *Buy Instagram Likes with Instant Delivery*, TWICSY, https://twicsy.com/buy-instagram-likes [https://perma.cc/V5K6-55JG] (last visited Mar. 12, 2022).

80. *See* Confessore et al., *supra* note 32; *see* Ellis, *supra* note 34.

81. *See generally* Confessore et al., *supra* note 32.

82. Ellis, *supra* note 34.

83. *See* Janith Weerasinghe et al., *The Pod People: Understanding Manipulation of Social Media Popularity via Reciprocity Abuse*, *in* WWW '20: PROCEEDINGS OF THE WEB CONF. 2020 1874, 1874 (2020) ("Pods are online groups designed to facilitate systematic *reciprocity abuse*, a term coined by DeKoven et al. describing an agreement between users to interact with each other's content, thereby increasing its popularity and consequent importance to the content curation algorithm.").

84. *See id.* at 1875.

return.[85] The primary selling point of pods is that they are usually free to participate in and are harder to detect than bot accounts.[86]

While pods may give a sense of realism that bot accounts do not, they are just as inauthentic.[87] The end result of both is the same: inflated follower counts and fake engagement.[88] Pods are just as much a business transaction as purchasing followers, and most pods even have rules or guidelines which actively state that members are not part of the pod to become friends but are simply there to serve as popularity generators for each other.[89]

Members of a pod may be real people, but they are likely not subject to an influencer's effects because they are not following or engaging with that influencer out of any care for what they have to say or emotional involvement with their content.[90]

3. Prevalence of Deception

It is estimated that over half of all social media influencers have utilized some form of fake influence during their careers.[91] Fake followers are commonplace at this point, and influencers try to downplay it as though it is an accepted industry practice. However, marketers and real followers do not share in that acceptance.[92] Real followers on social media websites have expressed an active disdain for influencers who use fake followers and engagement, and even though they are not losing money to the practice like brands are, real followers still dislike the practice.[93] A study in 2019 found

85. *See id.* at 1876 ("Some pods are designated for Instagram users who post about specific topics . . . .").

86. *See generally* Weerasinghe et al., *supra* note 83.

87. *See id.*; *see* John Boitnott, *How to Avoid Social Media Pods and Still Build an Audience*, ENTREPRENEUR (Oct. 12, 2020), https://www.entrepreneur.com/article/357164 [https://perma.cc/LWW4-TK5E] ("Although engagement pods aren't the same as 'buying' followers and likes, the process still essentially means you are creating fake likes and artificially enhanced engagement rates. The intent behind Instagram pods, for example, is to climb up the engagement ranks by manipulating Instagram's algorithm and follower counts, as opposed to organically targeting audience members that can convert into customers.").

88. *See id.*

89. *See* Emma Brown, *Do Instagram Pods Work? The Truth Behind Instagram's Latest Engagement Hack*, HOOTSUITE (Oct. 12, 2018), https://blog.hootsuite.com/instagram-pods/ [https://perma.cc/6VWC-5EZH] ("*Don't use the chat to chat* (this is purely business, no pleasantries allowed)."); *see generally* Weerasinghe et al., *supra* note 82, at 1874–76.

90. *See generally id.*

91. Eugene Tsaplin, *How to Avoid Getting Scammed by Influencers with Fake Followings*, ENTREPRENEUR (Jan. 21, 2022), https://www.entrepreneur.com/article/391195 [https://perma.cc/5ADX-GZ6G].

92. *See* Keith Weed, *When It Comes to Influencer Relationships, It's Complicated*, UNILEVER (July 22, 2018), https://www.unilever.com/news/news-search/2018/when-it-comes-to-influencer-relationships-its-complicated/ [https://perma.cc/G4N4-JMY9] (disavowing influencers who use fake followers or bots as dishonest in statement by the CMO of Unilever); s*ee* Fatih Cagatay Akyon & M. Esat Kalfaoglu, *Instagram Fake and Automated Account Detection*, *in* INNOVATIONS IN INTELLIGENT SYS. & APPLICATIONS CONF. 1, 1 (2019).

93. *See 71% Of Consumers Will Unfollow Influencers with Fake Followers*, SMART INSIGHTS (Oct. 22, 2019), https://www.smartinsights.com/online-pr/71-of-consumers-will-unfollow-influencers-with-fake-followers/ [https://perma.cc/4AL3-AG8P] [hereinafter *71% Of Consumers*].

that over 70% of U.S. and U.K. social media users would unfollow an influencer if they found out they had purchased influence at any point.[94] Brands and marketers do not take kindly to fake influence either.[95] Influencer marketing agency Mediakix ranked fake followers and engagement as the number one problem facing marketers, and 50% of marketers agreed it was their primary challenge on the job.[96] Marketers see fake followers as a form of lying or scamming.[97]

Fake activity subverts the goal of advertising: to reach real people, with real purchasing potential, via real engagement on an influencer's sponsored posts.[98] Brands are paying influencers to reach people whose opinions will be affected by the influencer's involvement and who may be willing to purchase a product after seeing an ad; they are not paying to receive likes from bots or random individuals who are only there as part of a business transaction.[99] Smaller influencers (and those hoping to become influencers) will pay bot accounts for follows and engagement to gain the attention of marketers, and influencers with more established followings or who are looking to grow their followings will purchase or trade for followers or engagement on their posts to move up into a new pay bracket.[100]

Today, an estimated one in ten accounts on Instagram is a bot account; a 2020 Twitter sweep of fake activity took out over 70 million bot accounts, and even that barely made a dent.[101] The problem touches engagement as well. Fifty percent of Instagram engagement with sponsored posts is estimated to be fake.[102] The practice of fake followers and engagement on social media

94. *See id.*

95. *See These Are the Influencer Marketing Trends Shaping 2021*, MEDIAKIX, https://web.archive.org/web/20220114085203/https://mediakix.com/influencer-marketing-resources/influencer-marketing-trends/ [https://perma.cc/567R-WUJF] (last visited Mar. 19, 2022); *see 95 Million Bots: One in Ten Instagram Accounts Is Fak*e, BASIC THINKING, https://www.basicthinking.com/bots-instagram-accounts-fake/ [https://perma.cc/ZDU3-CXED] (last visited Mar. 19, 2022) [hereinafter 95 Million Bots].

96. *See id.*

97. *See* Paquet-Clouston et al., *supra* note 77, at 5.

98. *See* Gian Fulgoni, *Fraud in Digital Advertising: A Multibillion-Dollar Black Hole*, J. ADVERT. RSCH., June 2016, at 122.

99. *See* Akyon & Kalfaoglu, *supra* note 92, at 1 ("The detection of fake engagement is crucial because it leads to loss of money for businesses, wrong audience targeting in advertising, wrong product predictions systems, and unhealthy social network environment.").

100. *See How to Spot Fake Followers on Instagram*, MEDIAKIX, https://web.archive.org/web/20210525120327/ttps://mediakix.com/blog/fake-followers-on-instagram/ [https://perma.cc/66VJ-RNWQ] (last visited Mar. 19, 2022).

101. *See 95 Million Bots*, *supra* note 95; Andrew Hutchinson, *New Fake Account Removals Highlight Twitter's Bot Problem Once Again*, SOC. MEDIA TODAY (Apr. 4, 2020), https://www.socialmediatoday.com/news/new-fake-account-removals-highlight-twitters-bot-problem-once-again/575488/ [https://perma.cc/TGS6-6V3R].

102. *See* Hutchinson, *supra* note 101; *see* Shareen Pathak, *Cheatsheet: What You Need to Know About Influencer Fraud*, DIGIDAY (Nov. 3, 2017), https://digiday.com/marketing/cheatsheet-need-know-influencer-fraud [https://perma.cc/G5FE-BRZ7].

platforms is prolific and only growing in popularity.[103] Searches and demand for fake followers went up by 71% in 2019 alone.[104] Most estimates place the losses generated by brands advertising to fake followers to be higher than a billion dollars per year.[105]

III. ANALYSIS

Something needs to be done to curb the rampant problem of fake activity on social media. Brands are losing, conservatively, over a billion dollars every year to influencers who think that fake followers are acceptable, and social media companies are financially disincentivized to help because their business model benefits from a higher user count generated by bots and the income generated by influencers.[106] With this kind of widespread loss and lack of assistance, we need to empower brands to take action against influencers who are abusing their business deals. Moreover, consumers feel deceived by influencers who purchase fake followers, and there is a vested public interest in seeing social media clean up its bot account problem and limit improper personal engagement.[107]

The Federal Trade Commission (FTC) has previously provided detailed guidance to influencers on when they need to disclose their involvement with or endorsement by a brand, but they have said far less on the practice of falsifying social media activity.[108] In the wake of the New York Times investigation and New York Attorney General's case against the bot-selling

103. *See generally 95 Million Bots*, *supra* note 95; s*ee* Pathak, *supra* note 102; *Why Are So Many People Still Buying Fake Social Media Followers?*, MEDIUM: GAIN (May 9, 2017), https://blog.markgrowth.com/why-are-so-many-people-still-buying-fake-social-media-followers-743d380b813b [https://perma.cc/QS7Q-EG53].

104. See *Demand for Fake Instagram Followers*, *supra* note 26.

105. *See* Anand et al., *supra* note 75, at 3; Megan Cerullo, *Influencer Marketing Fraud Will Cost Brands $1.3 Billion in 2019*, CBS NEWS, (July 25, 2019, 1:49 PM), https://www.cbsnews.com/news/influencer-marketing-fraud-costs-companies-1-3-billion/ [https://perma.cc/J7TP-NB8J].

106. *See* Brett Molina & Jessica Guynn, *Facebook Is Losing Users for the First Time Ever and Shares in Meta Have Fallen off a Cliff*, USA TODAY (Feb. 3, 2022, 7:00 PM), https://www.usatoday.com/story/money/2022/02/03/facebook-users-decline-meta-stock/6651329001/ [https://perma.cc/TF74-7ZX2] (discussing how for the first time ever, Facebook's user count dropped in 2022, and when it did, their stock plummeted.) We can infer from this that if Facebook made any real effort to remove the copious amount of bots on Instagram and Facebook, their user counts would drop, and their stock would tank again.

107. *See* Appel et al., *supra* note 21, at 89.

108. *See* FTC Press Release, *supra* note 32; *see generally* FTC, DISCLOSURES 101 FOR SOCIAL MEDIA INFLUENCERS, FTC ENDORSEMENT GUIDELINES (2019), https://www.ftc.gov/system/files/documents/plain-language/1001a-influencer-guide-508_1.pdf [https://perma.cc/J5TY-Y9CU].

website Devumi, the FTC did file a complaint against Devumi.[109] The case, which was the first of its kind, settled.[110] In a statement, the FTC said that:

> By selling and distributing fake indicators of social media influence to users of various social media platforms, the FTC alleges the defendants provided their customers with the means and instrumentalities to commit deceptive acts or practices, which is itself a deceptive act or practice in violation of the FTC Act.[111]

While this signals a willingness to tackle the practice of buying and selling fake followers, the FTC has yet to take any steps to enforce this against influencers themselves, nor has the FTC ever made mention of the practice of participating in social media pods.[112] Moreover, given that this complaint took place in 2019 and damages in the billions are still being racked up today, it seems apparent that something more needs to be done; this statement does, however, provide definitional guidance for holding those influencers accountable by other means.[113]

These influencers could likely be held liable under current fraud law if current definitions of fraud were interpreted in ways favorable to brands. Cases of influencers committing fraud are unlikely to be regularly brought in federal court, regardless of statute interpretation, for a number of reasons including: the comparatively small amount of money at issue for each individual sponsored post, federal statutes more frequently applying to fraud where the government is the victim or vehicle of the fraud, and actions already being undertaken within the industry in some states.[114] Thus any action within the industry will likely come at the state level.[115] There is some variation from state to state regarding the actual text of their fraud tests, but the applications of these tests are sufficiently similar that the outcome in one state regarding

109. *See id.*

110. *See* FTC Press Release, *supra* note 32; s*ee generally* Complaint for Permanent Injunction and Other Equitable Relief, FTC v. Devumi, LLC, No. 9:19cv81419 (S.D. Fla. Oct. 18, 2019); Athena Jones, *First on CNN: NY Attorney General Targets Fake Social Media Activity*, CNN (Jan. 30, 2019, 4:03 PM), https://www.cnn.com/2019/01/30/tech/new-york-attorney-general-social-media/index.html [https://perma.cc/M5B9-YWYJ].

111. *See* FTC Press Release, *supra* note 32.

112. *See id.*

113. *See id.*; *see* Ellis, *supra* note 34.

114. *See generally* 18 U.S.C. § 1341; *see* Jones, *supra* note 110; *see State vs. Federal Fraud Charges*, PRICE BENOWITZ LLP, https://criminallawyerwashingtondc.com/blog/state-vs-federal-fraud-charges/ [https://perma.cc/BT6H-FSG9] (last visited June 2022) ("The determination of whether a particular fraud case will be brought in Superior Court or federal court depends mainly upon whether or not the fraudulent conduct is in violation of federal law or involved an attempt to gain benefits through either a federal agency or federal program."); *see Federal Fraud Charges*, SPODEK L. GRP. (May 26, 2020), https://www.federallawyers.com/federal-fraud-charges/ [https://perma.cc/8EMT-HV5C] (listing types of federal fraud crimes such as mail fraud, wire fraud, Medicare fraud, and tax fraud which all involve the federal government as the vehicle or victim).

115. *See generally State vs. Federal Fraud Charges*, *supra* note 114.

the issue of influencer's fake followers will likely be the same in most others.[116]

The requirements for fraud in New York and California are applied so similarly that, when a conflict arises regarding which state's law should be applied to a fraud case, courts have found it unnecessary to explicitly decide which to apply.[117] The elements of a fraud claim in New York and California are: (1) a misrepresentation, (2) knowledge of falsity, (3) intent to defraud, (4) justifiable reliance, and (5) resulting damage."[118]

A. A Misrepresentation

The first element of liability for fraud is a misrepresentation of a material fact.[119] In California and New York, this element of fraud is generally interpreted as including any false representations, misrepresentations, some non-disclosures, and some concealments.[120] Many other states include non-disclosures and concealments in their definitions of fraud, but states have widely varying definitions of what types of non-disclosures and concealments are allowed to be considered fraudulent.[121] It is best, in determining the broader scope of potential liability of influencers for their social media

---

116. *Compare, e.g.*, Lazar v. Superior Court, 909 P.2d 981, 984 (Cal. 1996) ("The elements of fraud, which give rise to the tort action for deceit, are (a) misrepresentation (false representation, concealment, or nondisclosure); (b) knowledge of falsity (or 'scienter'); (c) intent to defraud, i.e., to induce reliance; (d) justifiable reliance; and (e) resulting damage." (quoting 5 WITKIN, SUMMARY OF CALIFORNIA LAW § 676 (9th ed. 1988))), *with* Spies v. Deloach Brokerage, Inc., 169 F. Supp. 3d 1365, 1374 (S.D. Ga. 2016) ("The tort of fraud consists of the following five elements: '(1) false representation or omission of a material fact; (2) scienter; (3) intention to induce the party claiming fraud to act or refrain from acting; (4) justifiable reliance; (5) damages.'" (quoting Lehman v. Keller, 677 S.E.2d 415, 417-18 (Ga. Ct. App. 2009))); *and* Bulbman, Inc. v. Nevada Bell, 825 P.2d 588, 592 (Nev. 1992) ("The[] elements [of fraud] are: 1. A false representation made by the defendant; 2. Defendant's knowledge or belief that the representation was false (or an insufficient basis for making the representation); 3. Defendant's intention to induce the plaintiff to act or to refrain from acting in reliance upon the misrepresentation; 4. Plaintiff's justifiable reliance upon the misrepresentation; and 5. Damage to the plaintiff resulting from such reliance." (citing Lubbe v. Barba, 540 P.2d 115, 117 (Nev. 1975))) (comparing all three definitions which each have different wording). There may be specific cases that would be affected by these differences, but the broader issue at play does not likely lend itself to any varying interpretation in these tests. Therefore, it is sufficient to address one.

117. *In re* Decade, S.A.C., LLC, 612 B.R. 24, 37 (Bankr. D. Del. 2020) ("With respect to fraud claims generally, New York and California define fraud using the same elements . . . . Additionally, the Parties similarly define fraud in the execution, fraudulent inducement, and fraudulent misrepresentation.").

118. *Id.* at 24.

119. Lazar v. Superior Court, 909 P.2d 981, 984 (Cal. 1996).

120. Petersen v. Allstate Indem. Co., 281 F.R.D. 413, 419 (C.D. Cal. 2012) ("[M]isrepresentation . . . includes either false representation, concealment or nondisclosure . . . .") (internal quotations omitted); *In re* Decade, S.A.C., LLC, 612 B.R. 24, 37 (Bankr. D. Del. 2020); Peter R.J. Thompson, *An Outline of 23 California Common Law Business Torts*, 13 PAC. L.J. 1, 6 (1981).

121. *See* McCullough v. World Wrestling Ent., Inc., 172 F. Supp. 3d 528, 563 (D. Conn. 2016) (laying out where Connecticut courts have a different definition and different application of non-disclosed facts in a fraud case).

activities, to stick to the broadest applicable portions of that definition and consider just false representations and misrepresentations of material fact.

The use of bot accounts for fake followers and engagement would almost certainly be a false representation of a material fact.[122] Material facts are defined "to mean that, counterfactually, the plaintiff would have acted differently but for the alleged misrepresentation or omission."[123] Followers and engagement are the driving forces behind why an influencer is chosen to market a brand, and they are the determining factor in what an influencer will get paid for a given piece of content.[124] Influencers have used bot accounts to falsely represent this number to brands.[125] Brands have no use or wish to market to bot accounts because the purpose of advertising is to convince real people to purchase a product or service.[126] Bots are not real people and do not have a disposable income to spend on a brand, so the brand does not care for them to see their ads.[127] Because they know no brand would willingly pay for a bot to see their account, influencers falsely represent these bots as real people with the hope of getting paid extra for the fake influence as well as their real influence.[128] But for the false representation of their follower and engagement counts, an influencer's compensation would be drastically different, presuming the brand would work with them at all; this is, therefore, a false representation of a material fact.[129]

This issue includes an added layer of complexity in the realm of social media pods. Follower and engagement counts are still material facts, but some influencers may argue that engaging in pods is not a false or misleading representation.[130] Pods are made up of real people and real engagement, even if done so for their members' own benefit.[131] If a follower or engagement count is merely a representation of how many people are present on an influencer's page to view and interact with an ad, then it could be said that the members of the pod meet this very low qualification of being "real." Moreover, some pods require members to share some common trait which would have them fall within the target market of a brand.[132] It also cannot be assumed that, simply because members of a pod are getting something out of engaging with an influencer's content, that they are not affected by ads they

122. *See generally* Paquet-Clouston et al., *supra* note 77, at 5 (indicating how marketers already see this as fraud and misrepresentation).

123. FDIC v. Murex LLC, 500 F. Supp. 3d 76, 111 (S.D.N.Y. 2020) (quoting City Trading Fund v. Nye, 72 N.Y.S.3d 371, 378 (Sup. Ct. 2018)).

124. *See* Paquet-Clouston et al., *supra* note 77.

125. *See* Appel et al., *supra* note 21.

126. *See id.*

127. *See* Harry Kabadaian, *How Bots Steal Your Online Advertising Budget*, ENTREPRENEUR (July 13, 2018) https://www.entrepreneur.com/article/313943 [https://perma.cc/Y672-5YB7].

128. *See id.*

129. HENNESSY, *supra* note 73.

130. *See* Lauren O'Neill, *Influencers Have Secret Engagement Pods. I Joined One*., VICE (Nov. 9, 2021, 4:30 AM), https://www.vice.com/en/article/pkpy5g/influencers-have-secret-engagement-farms-i-joined-one [https://perma.cc/MD6B-FGLW].

131. *See* Weerasinghe et al., *supra* note 83, at 1874.

132. *See id.* at 1877–78.

interact with. Ads often unintentionally affect consumers.[133] Just because members of a pod are there for their own benefit does not mean they will not, for example, think a product on a fellow pod member's page is intriguing.[134] The only way for a brand to be certain that influencer pods would be considered false representation would be to specify when dealing with influencers that they are paying them based on their number of followers, excluding pod activity. In this scenario, if an influencer fails to accurately convey the prevalence of pod activity on their account, brands could argue that this constitutes a material misrepresentation.

B. Knowledge of Falsity

The second element of liability for fraud, which is easily settled in this scenario, is knowledge of falsity.[135] Knowledge of falsity is a condition which can be met in several ways; the primary definition relevant for influencers would be: "knows or believes that the matter is not as he represents it to be … or knows that he does not have the basis for his representation that he states or implies."[136] When individuals purchase followers or become involved in pod activity, they assuredly have knowledge of their own actions.[137] They are aware of how many fake followers and how much fake engagement they have paid for. They also know how many members are in their pod following and engaging with their activity in exchange.[138] Thus, they have specific knowledge of how much they have falsely represented their influence.

Knowledge of falsity does, however, present one potential defense for those who purchased from bot accounts. In the FTC's Devumi complaint, several of Devumi's customers did not know that they were paying for fake activity; rather, they thought they were paying for authentic endorsements.[139] This would suggest that some Devumi users thought they were paying real people for the likes and follows they received, a legitimate misconception

133. *See generally* Jenna Gross, *The Subconscious Implications of Marketing*, FORBES: AGENCY COUNCIL (Dec. 19, 2017, 8:30 AM), https://www.forbes.com/sites/forbesagencycouncil/2017/12/19/the-subconscious-implications-of-marketing/ [https://perma.cc/WX4B-HVAX] (explaining the subconscious effects of ads). Some, though certainly not all, will affect the subconscious of consumers around them in some way; particularly where you can be guaranteed that consumers are actually looking at the ad.

134. *See, e.g.*, ZOE GANNON & NEIL LAWSON, THE ADVERTISING EFFECT: HOW DO WE GET THE BALANCE OF ADVERTISING RIGHT, 8 (2010), https://www.compassonline.org.uk/wp-content/uploads/2013/05/The-advertising-effect-compass.pdf [https://perma.cc/7NQ4-BDLK].

135. Gold v. L.A. Democratic League, 122 Cal. Rptr. 732, 743 (Ct. App. 1975); Lazar v. Superior Court, 909 P.2d 981, 984 (Cal. 1996).

136. Cummings v. HPG Int'l, Inc., 244 F.3d 16, 25 (1st Cir. 2001).

137. *See* Weerasinghe et al., *supra* note 83, at 1874 (showing how influencers get involved in pods and that it is a willing process).

138. *See id.* at 1877–78.

139. *See* FTC Press Release, *supra* note 32.

given that bots are designed to look as much like real people as possible.[140] Whether they genuinely had this misconception or not, users of the website could easily make this claim because many bots impersonate real people or take other measures to look like legitimate accounts to avoid being caught in sweeps by social media websites.[141] However, it seems unlikely that the FTC intended to give influencers a sweeping defense to fraud claims where they merely needed to claim they reasonably thought the bots were real people. In the same complaint, the FTC implied that purchasing fake followers was deceptive, which supports the idea that they did not intend to create a defense.[142] If Devumi users operated under a reasonable belief that the followers they had purchased were real people, this would still put their activity on par with pods in that they believed the followers were real people but knew they were not following them because of a genuine affinity toward the influencer.

This element also protects honest influencers from overeager brands attempting to catch them up in lawsuits. Bot accounts in particular are known to follow people who have not paid for their services in order to create a veneer of legitimacy as a defense against being taken down.[143] Most public individuals have at least some following from bot accounts as a result of this activity.[144] Even honest influencers who have never paid for bot accounts will likely have some number of bot followers simply due to their reach.[145] They do not have knowledge of particular bot accounts or how many bot accounts follow them, so they have no ability to disclose that information to brands in any way that would suggest knowledge of the falsity of those accounts.

C. Intent to Defraud

Intent is the next element of a fraud claim.[146] California courts have held that "it is the element of fraudulent intent or intent to deceive that distinguishes it from actionable negligent misrepresentation and from nonactionable innocent misrepresentation."[147] It is the element of intent which makes fraud actionable, irrespective of any contractual or fiduciary duty one

140. Asaf Greiner, *The Hidden Costs of Identity Theft*, FORBES (June 1, 2018, 7:30 AM), https://www.forbes.com/sites/forbesagencycouncil/2018/06/01/the-hidden-costs-of-identity-theft/ [https://perma.cc/PEU2-K95X].

141. Adrianne Jeffries, *It's Your Face. It's Your Photos. Meet the Creepiest Kind of Instagram Spambot*., VERGE (Sept. 3, 2014, 8:29 AM), https://www.theverge.com/2014/9/3/6097891/its-your-face-its-your-photos-meet-the-new-creepy-breed-of-instagram-spambot [https://perma.cc/V6V2-RCAJ].

142. *See* FTC Press Release, *supra* note 32 ("By selling and distributing fake indicators of social media influence to users of various social media platforms, the FTC alleges the defendants provided their customers with the means and instrumentalities to commit deceptive acts.").

143. Kate Moffatt, *This Is Why You Keep Getting Those Random Instagram Followers*, ELLE AUSTL. (June 24, 2017, 11:39 PM), https://www.elle.com.au/culture/random-instagram-followers-13543 [https://perma.cc/4YNP-8CA6].

144. *See id.*

145. *See id.*

146. *In re* Decade, S.A.C., LLC, 612 B.R. 24, 37 (Bankr. D. Del. 2020).

147. City of Atascadero v. Merrill Lynch, 80 Cal. Rptr. 2d 329, 355 (Ct. App. 1998).

party might owe another."[148] Under California law, this intent must be to deceive, not merely to induce.[149]

1. Intent with Fake Engagement

Intent is a unique issue, which is best divided not along the lines of pod or bot accounts, but on whether the fraudulent activity occurred through engagement or followers. While fake engagement can be, and often is, directed at non-sponsored posts in the hopes of making the influencer seem popular enough to warrant partnership with a brand and to hide the use of fake engagement on brand posts, fake engagement on non-sponsored posts is an issue more readily dealt with by false inducement claims than fraud.[150] When a sponsored post is the subject of the fake engagement, it usually makes the intent element clear. Fake engagement, when purchased through bot accounts, is purchased for a particular post sponsored by a particular brand.[151] Bots do not hand out engagement for free, nor do they randomly distribute likes to various posts on an account.[152] In pods, fake engagement only occurs on posts created at the pod's designated time to reciprocate activity or by linking to a particular post in the pod's chat.[153] Either way, fake engagement does not happen on every post and is a specifically-directed activity.[154] Thus, when fake engagement occurs on a sponsored post, it is specific, and its intent to gain additional revenue for the influencer is evident in the act.[155]

2. Intent with Fake Followers

The difficulty in proving intent in the case of fake followers on social media is that, in many cases, the followers are obtained at a different time

148. *See id.*

149. *See* Sun 'n Sand, Inc. v. United Cal. Bank, 582 P.2d 920, 942 (1978) (holding that it was not sufficient that the bank had presented checks, as that was merely inducement, but they also needed an intent to deceive Sun n' Sand, which was not present); *see* Thompson, *supra* note 120, at 7.

150. *See generally* Weerasinghe, *supra* note 83

151. *See generally Buy Instagram Likes with Instant Delivery*, BUZZOID, https://buzzoid.com/buy-instagram-likes/ [https://perma.cc/NLZ4-KTLN] (last visited Apr. 10, 2022) [hereinafter *Buy Instagram Likes*] (showing in the 'What information do I need to provide' section, Buzzoid's website expressly says "[t]he only information we need is your username and instructions regarding which photo or video you want to receive the likes," meaning the purchaser has to choose and instruct Buzzoid exactly which post to like). *Id.*

152. *See id.*

153. *See* Weerasinghe et al., *supra* note 83, at 1877–78; *see* Boitnott, *supra* note 86.

154. *See* Weerasinghe et al., *supra* note 83.

155. *See generally* Pathak, *supra* note 103 (documenting that 50% of engagement on sponsored posts in a day was fake.) This is an extremely high ratio of inauthentic activity, and given that influencers have to go out of their way to purchase fake activity and profit off of engagement on sponsored posts, there can be an inferred relationship between that profit and the purchase of these followers.

than the agreement with the brand is made.[156] In pods, it is slightly easier than with bot accounts to find fraudulent intent regarding followers because, regardless of how long ago an individual chose to join the pod, pods require continuous community involvement.[157] Real people will likely not continue to follow an individual that they only followed out of reciprocity if that individual ceases to benefit them in return.[158] If an influencer decides to part ways with a pod or cease their pod activity, they will likely be unfollowed by a majority of the pod members and could easily argue that pod members who do not unfollow them when they cease their pod activity are now passive, willing followers because they are receiving nothing in return.[159] Continuous involvement in a pod is evidence of the intent to defraud.

Bot accounts, however, are not a continuous action; they are a single event purchase.[160] If that purchase happens near striking a deal with a marketer, then intent can be inferred from the act. For example, if an influencer purchases followers shortly before opening negotiations with a brand, then the intentional link between the act of purchase and deceiving the brand is obvious. It is less obvious when an influencer purchases followers well in advance; influencers often purchase followers early in their careers to help build an audience or simply because they want to be perceived as popular.[161] At the time of purchase, these followers are not being used to deceive a brand, and the influencer may not have considered using them in a marketing deal. However, the influencer is unlikely to admit to the purchase, even if they do have the knowledge. While social media websites will occasionally attempt to purge bots, as previously referenced, the vast majority of fake accounts, even when reported, will not be removed by a social media website.[162] Failure to disclose the prior purchase during brand negotiations is evidence of the intent to profit from the deception, even if the original purchase was not made specifically to deceive a brand.

D. Justifiable Reliance

The fourth element of a fraud claim is a justifiable reliance on the fraudulent information (in this case, follower count and engagement rates)

156. *See generally Buy Instagram Likes*, *supra* note 151 (stating that the process by which fake likes are purchased requires directing them to a specific post, but the process by which fake followers are purchased only indicates what account you want them on and what quality of fakes you want, and as Buzzoid states, premium followers will last for years before Instagram gets around to taking them down).

157. *See* Weerasinghe et al., *supra* note 83; *see* Boitnott, *supra* note 87.

158. *See* Weerasinghe et al., *supra* note 83; s*ee* Boitnott, *supra* note 87.

159. *See* Weerasinghe et al., *supra* note 83; s*ee* Boitnott, *supra* note 87.

160. *See generally Buy Instagram Likes*, *supra* note 151.

161. *See id.* (discussing what package of followers to purchase and considering what stage in the process of becoming an influencer the account followers are being purchased for is at, as newer accounts need premium followers that they are less likely to lose and are more likely to hold up to scrutiny given how few followers they have).

162. Cox, *supra* note 26; Maya Kosoff, *Can Twitter Purge Its Bots Without Killing Its Bottom Line?*, VANITY FAIR (June 27, 2018), https://www.vanityfair.com/news/2018/06/can-jack-dorsey-twitter-purge-bots-without-killing-bottom-line [https://perma.cc/7L79-ZJDG].

that has been given to the brand in question.[163] While it would be reasonable for a normal person to rely on publicly available numbers to tell them how many followers and how much engagement another individual has, brands are different. They have greater resources and knowledge at their disposal which may significantly affect what level of reliance is justifiable on their part.

There are multiple instances in which it would be unjustifiable for a brand to rely on information given to them by an influencer. One example would be if a brand investigated an influencer to determine whether their reported followers and engagement are real. The fact that a brand independently investigated the information given to it makes it difficult to claim it relied on the falsity.[164] Another example would be if the brand's experience and intelligence should counsel against reliance.[165] Some brands have a large amount of experience at their disposal which may negate their justifiable reliance on influencer's followers and engagement.

1. Investigations

Brands primarily rely on bot detection websites to conduct independent investigations of an influencer's follower count.[166] These websites are readily available across the Internet, and most are free for anyone to use.[167] However, the accuracy of the majority of these websites is questionable at best, and they tend to overestimate the number of bot accounts associated with an influencer.[168] Many will catch inactive accounts who influencers have no control over just as readily as they will catch bots.[169] Even methods that claim to be more advanced tend to be newer versions of the same technology, utilizing publicly viewable features of an account to label them bot or human which can be avoided by sophisticated bots and the use of pods.[170]

---

163. Gold v. L.A. Democratic League, 122 Cal. Rptr. 732, 739 (Ct. App. 1975); Lazar v. Superior Court, 909 P.2d 981, 984 (Cal. 1996).

164. *See* Outdoor Cent., Inc. v. GreatLodge.com, Inc., 688 F.3d 938, 942 (8th Cir. 2012).

165. Scottish Heritable Tr., PLC v. Peat Marwick Main & Co., 81 F.3d 606, 615 (5th Cir. 1996) ("The justifiableness of the reliance is judged in light of the plaintiff's intelligence and experience.").

166. *Fake Influencer & Credibility Tool*, GRIN, https://grin.co/fake-influencer-tool/ [https://perma.cc/62U6-AN2C] (last visited Apr. 10, 2022); FAKECHECK.CO, https://www.fakecheck.co/ [https://perma.cc/KQ6U-5MQY] (last visited Apr. 10, 2022).

167. *See Fake Influencer & Credibility Tool*, *supra* note 166; see FAKECHECK.CO, *supra* note 166.

168. *See* James Parsons, *How Accurate Is the Twitter Audit Follower Checker?*, FOLLOWS.COM (Feb. 12, 2022), https://follows.com/blog/2022/02/accurate-twitter-audit-checker [https://perma.cc/3CEV-SBVY] (discussing the Twitter Audit, an app that is an excellent example of how inaccurate these checking websites and apps can be because these programs cannot distinguish between bots and the inactive or low effort accounts of real people).

169. *See id.*

170. *See generally* Shad Mohammad et al., *Bot Detection Using a Single Post on Social Media*, *in* 2019 THIRD WORLD CONF. ON SMART TRENDS IN SYS. SEC. & SUSTAINABILITY 215 (2019) (mentioning previous methods of detection which earmark accounts based on viewable information, but also suggesting a new method of far more advanced detection which has yet to be widely implemented).

There is little case law that suggests that brands, or others alleging fraud, would have an active responsibility to conduct independent investigations into influencer metrics.[171] Brands are welcome to choose whether to investigate information or rely on it, but if they choose to investigate, they may not be able to claim justifiable reliance.[172] Use of these websites may be argued by influencers to be a form of independent investigation, which makes brands aware of their fake followers and thereby makes reliance unjustifiable.[173] The inaccuracy of these methods would be a brand's best defense in instances where they did use one. These websites were not created in partnership with the social media platform itself; they are developed by third parties who have varying degrees of expertise and credibility.[174] Even if brands choose to utilize them, they would not be able to rely on them for any degree of accuracy.[175] If they did, due to the detection of inactive users in most cases, brands would likely underpay influencers, which causes an entirely different problem.

Independent investigation is a less pressing issue for brands regarding pods. There are many pods scattered across the Internet on a wide variety of platforms, and there is no way for a brand to be sure they have found and checked them all for involvement with a given influencer, even if they were to attempt conducting an independent investigation.[176] If pods were open to the public, with an obvious membership, it would defeat their entire purpose of seeming to offer each other organic support. Such investigations, even when attempted, could not be considered thorough or confirmatory in any way.[177]

2. Brand Experience

The experience of brands on social media is a more significant defense that may be put forward by influencers, though it varies greatly depending on the brand. Many brands have a long and storied history involving influencers

171. Field v. Mans, 516 U.S. 59, 70 (1995) (commenting on the Restatement's expounding upon of justifiable reliance by explaining that a person is justified in relying on a representation of fact "although he might have ascertained the falsity of the representation had he made an investigation." (quoting RESTATEMENT (SECOND) OF TORTS § 540 (AM. L. INST. 1977))).

172. *See generally id.*

173. *See id.*

174. FAKECHECK.CO, *supra* note 166.

175. *See* Parsons, *supra* note 168.

176. *See* Weerasinghe et al., *supra* note 83.

177. *See id.*

and celebrities in their marketing campaigns.[178] It is difficult, given the prevalence of fake activity on social media, to believe that any brand reliant on influencers has *not* been affected by fake activity of some kind and is not acutely aware of the dangers. If we are to believe that one in ten accounts on Instagram is fake and that Facebook leaves up 95% of reported bot accounts, influencers may argue that fake activity is generally so common as to be an expected part of social media.[179] Influencers, therefore, would not be wrong to suggest that influencer-savvy brands should anticipate this to some degree. Even within those parameters, influencer-savvy brands would still have some room for argument depending on the particular facts of an influencer's case, such as just how much falsification a brand should expect from a given influencer and how open they were with influencers in asking about their fake activity.

However, the fact that some brands have sufficient expertise that they cannot claim to reasonably rely on influencers does not mean that all brands do. Courts consider a party's personal expertise to determine whether it reasonably relied on the fraudulent activity; they do not suggest that merely because one party possesses sufficient knowledge, a similarly situated party would be expected to as well.[180]

Because of trends in advertising, many businesses who lack resources and are technologically unaware have been pushed into the influencer space with little understanding of what they are getting themselves into. Moreover, small business relationships with influencers are often initiated by the influencers themselves.[181] Influencers asking for free stuff from small business owners is a common exploitative trend in modern social media, and the Internet is littered with examples of small businesses being pressured or blackmailed into giving influencers their services for free.[182] These businesses are often uninformed of the risks and consequences associated with social media influencers and sometimes do not even mean to or want to be involved in the first place.[183] They are relying on influencers out of necessity and are placing their trust in the influencers as individuals, hoping that they will not be led astray.[184]

---

178. *See generally* Peter Suciu, *History of Influencer Marketing Predates Social Media by Centuries – But Is There Enough Transparency in the 21st Century?*, FORBES (Dec. 7, 2020, 9:43 AM), https://www.forbes.com/sites/petersuciu/2020/12/07/history-of-influencer-marketing-predates-social-media-by-centuries--but-is-there-enough-transparency-in-the-21st-century/?sh=21f0d09340d7 [https://perma.cc/WAG2-5JDQ]; *see, e.g.*, Naomi Fry, *Fake Famous and the Tedium of Influencer Culture*, NEW YORKER (Feb. 20, 2021), https://www.newyorker.com/culture/on-television/fake-famous-and-the-tedium-of-influencer-culture [https://perma.cc/2CGB-GP8Q] (discussing the HBO documentary "Fake Famous," which promotes buying fake followers as a means of getting famous).

179. Cox, *supra* note 26.

180. *See* Field v. Mans, 516 U.S. 59, 70 (1995).

181. Hannah Dobrogosz, *28 Screenshots That Prove Influencers Are Absolutely Out of Control*, BUZZFEED (Dec. 13, 2021), https://www.buzzfeed.com/hannahdobro/entitled-influencers [https://perma.cc/H66L-3YQ3].

182. *See id.*

183. *See generally* Akyon, *supra* note 92 (explaining the risks with influencers and how businesses are unaware when they initially pay on).

184. *See* Carr, *supra* note 40.

There is one category of brands, however, to whom influencers can say, categorically, that they owe no reliance, and that is those who actively contribute to the problem. In recent years, some brands have even been known to engage in the purchase of fake followers themselves.[185] It would be difficult for such brands to claim that they expected an influencer's follower or engagement count to be accurate given their own unclean hands. In those cases, fake followers are a two-way street.[186] For all the brand knows, they are paying an influencer more than they are worth; for all the influencer knows, they are being contracted to work for a company who is not nearly prominent enough to warrant their partnership.

Purchasing followers makes some brands active contributors to the problem. To assume that the mere existence of fake followers necessarily creates an atmosphere of unreasonable reliance would be to ignore the hard work of thousands of influencers who have built their popularity honestly. In instances where it can be proven that brands themselves purchased undisclosed followers, influencers would easily be able to negate a justifiable reliance, though brands who have not themselves engaged in fraudulent activity would still have recourse.

E. Damages

There is a clear monetary damage to brands in these instances: brands are overpaying influencers by an amount equal to their faked activity.[187] Calculating what an influencer would be worth based on real engagement and real followers can determine how much a brand overpaid, and thereby how much they are owed for the fraudulent activity. Moreover, brands may wish to argue a more arbitrary level of damage experienced to their brand's name by being associated with fraud.[188] A brand's reputation may or may not experience a negative impact if it becomes clear that they were associated with influencers who fake social media influence.[189] Real followers disapprove of fake activity, and will unfollow influencers whose reliance on fake activity has been made public.[190] The backlash may extend to the influencer's affiliated brands as well. Moreover, depending on the company, some brands may simply look bad by not being aware of the fraud,

185. *See* Nicole Perlroth, *Researchers Call Out Twitter Celebrities with Suspicious Followings*, N.Y. TIMES: BITS (Apr. 25, 2013, 4:57 PM), https://bits.blogs.nytimes.com/2013/04/25/researchers-call-out-twitter-celebrities-with-suspicious-followings/ [https://perma.cc/2Q9Z-BA5F] (accusing brands like Pepsi of having fake followings).

186. *See id.*

187. *See* Santora, *supra* note 69.

188. Masa Mustafa Al-Qatami, The Effects of Social Media Influencer Attributes on Collaborating Brand Credibility and Advocacy (2019) (M.S. thesis, Qatar University, College of Business and Economics) (available at https://qspace.qu.edu.qa/bitstream/handle/10576/11689/Masa%20Al-Qatami_OGS%20Approved%20Thesis.pdf?sequence=1&isAllowed=y) [https://perma.cc/3SCG-T6V4].

189. *See id.*

190. *See 71% Of Consumers*, *supra* note 93.

particularly if their brand has a reputation for honesty or integrity. These damages are not as easily quantifiable and would depend on specific facts.

IV. CONCLUSION

Influencer marketing is a growing industry with a growing problem. In the face of social media companies who refuse to take action, brands need to be empowered to clean up social media platforms that are now serving as their place of business. Holding influencers accountable for the fraud they are perpetrating against brands would not only curb the monetary damages brands experience but would also disincentivize bot account creators from continuing to flood social media with inauthentic behavior. Being an influencer is a difficult job, and those who are unwilling to put in the work to earn authentic growth should not be allowed to jump to the top of their industry simply because that industry exists online.

We Know What's in Your Wallet: Data Privacy Risks of a Central Bank Digital Currency

Thompson J. Hangen*

TABLE OF CONTENTS

* J.D., May 2023, The George Washington University Law School; Senior Notes Editor, *Federal Communications Law Journal*, Volume 75; B.A., May 2017, Russian and Post-Soviet Studies, The College of William & Mary. I would like to thank Meredith Rose and Natasha Nerenberg for their encouragement, feedback, and support throughout the writing process. I would also like to thank Steve Young for getting me started on learning about blockchain and Matt Gertler for providing opportunities to further develop knowledge in legal issues with blockchain and cryptocurrencies. Finally, I would like to thank my wife for her constant support throughout my law school journey.

I. INTRODUCTION

Governments worldwide are interested in developing and issuing digital currencies, also known as central bank digital currencies ("CBDC," or plural, "CBDCs").[1] A CBDC is issued by a central bank using technology similar to cryptocurrencies[2] and is legal tender.[3] Issuing a CBDC may give governments additional powerful monetary policy tools,[4] but because of the technology involved, also allows governmental agencies to collect massive amounts of identifiable financial data.[5] Where consumer data is collected, consumer data should be protected; when that collection includes every single system transaction, data must be all the more strictly guarded.[6]

The Federal Reserve System ("FRS"), which operates as the central bank in the United States, is responsible for "conducting the nation's monetary policy" and "promoting consumer protection."[7] Implementation of a central bank digital currency in the United States would provide additional policy levers to conduct monetary policy but would also extend the role of the FRS from "promotion" of consumer protection to active collection of consumer data at an unprecedented level.[8] This consumer data would connect an individual to every single financial transaction they, or others connected to

---

1. *See Central Bank Digital Currency Tracker*, ATLANTIC COUNCIL, https://www.atlanticcouncil.org/cbdctracker/ [https://perma.cc/V5XW-2AA7] (last visited Nov. 17, 2021) (tracking development of CBDCs across 90 countries); Turner Wright, *IMF Director: 110 Countries Are 'At Some Stage' of CBDC Development*, COINTELEGRAPH (Oct. 5, 2021), https://cointelegraph.com/news/imf-managing-director-110-countries-are-at-some-stage-of-cbdc-development [https://perma.cc/7YPJ-E7V2].

2. *CBDC vs Cryptocurrency: What Are the Core Differences?*, SHRIMPY ACAD. (May 20, 2021), https://academy.shrimpy.io/post/cbdc-vs-cryptocurrency-what-are-the-core-differences [https://perma.cc/8SCT-6FY6] [hereinafter *CBDC vs Cryptocurrency*].

3. *See* Matthew Green & Peter Van Valkenburgh, *Without Privacy, Do We Really Want a Digital Dollar?*, COIN CTR. (Apr. 30, 2020), https://www.coincenter.org/without-privacy-do-we-really-want-a-digital-dollar/ [https://perma.cc/PAP6-LQN2]. *Contra* Anatoly Kurmanaev et al., *Bitcoin Preaches Financial Liberty. A Strongman Is Testing That Promise*, N.Y. TIMES (Oct. 12, 2021), https://www.nytimes.com/2021/10/07/world/americas/bitcoin-el-salvador-bukele.html [https://perma.cc/83FW-X9DS] (describing how El Salvador has made Bitcoin—a cryptocurrency—legal tender).

4. *See* Brandon Van Niekerk, *Central Bank Digital Currencies: A Technocratic Fallacy*, BITCOIN MAG. (Oct. 17, 2021), https://bitcoinmagazine.com/culture/central-bank-digital-currencies-bitcoin [https://perma.cc/YF4B-FW3E].

5. *See, e.g.*, Ajay S. Mookerjee, *What if Central Banks Issued Digital Currency?*, HARV. BUS. REV. (Oct. 15, 2021), https://hbr.org/2021/10/what-if-central-banks-issued-digital-currency [https://perma.cc/3XU6-8Z26] (discussing how China had collected information on over 500 million transactions by the end of September 2021—mere months after rolling out a limited pilot of a digital Yuan CBDC).

6. *See* Grp. of Seven [G7], *Public Policy Principles for Retail Central Bank Digital Currencies (CBDCs)*, at 7–8 (Oct. 14, 2021), https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1025235/G7_Public_Policy_Principles_for_Retail_CBDC_FINAL.pdf [https://perma.cc/D4AS-5J55] [hereinafter *Public Policy Principles*].

7. *About the Fed*, BD. GOVERNORS FED. RSRV. SYS., https://www.federalreserve.gov/aboutthefed.htm [https://perma.cc/F9Y2-7RRR] (last visited Nov. 22, 2021).

8. *See, e.g.*, *CBDC vs Cryptocurrency*, *supra* note 2.

them, make using a CBDC.[9] While this data could likely be anonymized, it could still be possible to connect an individual to data—including their demographic data, geographic location, financial transaction history, and even the types of personally identifiable information generally collected by banks today to open an account.[10] This massive amount of information carries unique risks for consumers if compromised.[11] Safe implementation of a CBDC requires both stringent legislation aimed at safeguarding consumer data and the institutional competence within the FRS necessary to realize such safeguards.

The Federal Reserve has indicated that development of a CBDC for the United States is a "priority."[12] While development and implementation could take years, appropriate data privacy protections must be built into the development of a CBDC from the outset.[13] Information from the FRS about how a CBDC might be implemented in the United States is still under discussion and may yet include some protection for consumer data.[14] However, adequate data privacy standards, discussed in Section IV below, will likely require an expansion of federal laws (such as the Gramm-Leach-Bliley Act) to cover the unique types of data collected as part of the routine functionality of a CBDC.[15] Such coverage would require Congress to explicitly expand the scope of the FRS.[16]

Interest in a central bank digital currency is not limited to the United States; as of May 2022, at least 87 countries, representing more than 90% of global GDP, are exploring, actively developing, or in the process of implementing a CBDC.[17] Seven countries have fully implemented a CBDC,

9. *See* Van Niekerk, *supra* note 4.

10. *See id.*

11. *See id.* (describing the risk to consumers from the collection of CBDC data as "a perfect honeypot for hackers, fraudsters and the corrupt").

12. Sarah Hansen, *Fed Chair Powell Says Digital Dollar Is a 'High Priority Project'*, FORBES (Feb. 23, 2021, 1:21 PM), https://www.forbes.com/sites/sarahhansen/2021/02/23/fed-chair-powell-says-digital-dollar-is-a-high-priority-project/ [https://perma.cc/6T2E-63BM].

13. *See* BD. OF GOVERNORS OF THE FED. RSRV. SYS, MONEY AND PAYMENTS: THE U.S. DOLLAR IN THE AGE OF DIGITAL TRANSFORMATION 13 (2022) [hereinafter MONEY AND PAYMENTS], https://www.federalreserve.gov/publications/files/money-and-payments-20220120.pdf [https://perma.cc/QUQ6-LLNB].

14. *See* Andrew Ackerman, *Fed Prepares to Launch Review of Possible Central Bank Digital Currency*, WALL ST. J. (Oct. 5, 2021, 5:30 AM), https://www.wsj.com/articles/fed-prepares-to-launch-review-of-possible-central-bank-digital-currency-11633339800 [https://perma.cc/C3L2-W5WU] (detailing how the Federal Reserve plans to publish a discussion paper on development and use of a CBDC in 2021); MONEY AND PAYMENTS, *supra* note 13, at 19–20 (minimally discussing data privacy concerns in issuance of a CBDC).

15. *See* Fara Soubouti, Note, *Data Privacy and the Financial Services Industry: A Federal Approach to Consumer Protection*, 24 N.C. BANKING INST. 527, 528 (2020) (discussing the need for an expansion of the Gramm-Leach-Bliley Act to include data that commercial banks *already* routinely collect).

16. *See* Christopher J. Waller, Member, Bd. of Governors of the Fed. Rsrv. Sys., CBDC - A Solution in Search of a Problem?, Speech at the American Enterprise Institute 2–3 (Aug. 5, 2021), https://www.bis.org/review/r210806a.pdf [https://perma.cc/T5Y2-CU3B].

17. *See Central Bank Digital Currency Tracker*, *supra* note 1; Wright, *supra* note 1.

and an additional seventeen are currently piloting one.[18] At the 2021 G7 Summit, member countries[19] released thirteen principles to which a CBDC should adhere, demonstrating the importance of near-term CBDC development to major world economies.[20] The G7 recognized that each nation's data privacy laws differ but agreed that generally, a CBDC "must protect the privacy of users, including by requiring that the processing of their personal data is subject to laws governing privacy and the collection, storage, safeguarding, disposal and use of personal data that are enforceable in the jurisdiction."[21] However, notwithstanding the emergence of CBDCs worldwide, there exists no comprehensive data privacy standards or guidelines that countries can use as a benchmark for consumer data protection.[22]

To address these interests, Congress should explicitly expand the scope of the FRS.[23] Section II.A of this Note provides a high-level discussion of technologies used to create blockchains and how they can be used to create a centralized ledger for central bank digital currencies, as well as a discussion of how CBDCs differ from more common cryptocurrencies. Section II.B considers the case for and against issuance of a CBDC, including data privacy concerns. Section III reviews how federal financial data privacy laws (especially the Gramm-Leach-Bliley Act) currently provide for consumer data protection and storage, and the extent to which such laws might cover CBDC-related data.[24] Section IV urges Congress to enact a unified data privacy standard that encompasses CBDC data and to empower the Federal Reserve System to collect, safely store, and protect consumer data. Section V

18. *See Central Bank Digital Currency Tracker*, *supra* note 1; Jinia Shawdagor, *Asian CBDC Projects: What Are They Doing Now?*, COINTELEGRAPH (Oct. 16, 2021), https://cointelegraph.com/news/asian-cbdc-projects-what-are-they-doing-now [https://perma.cc/E276-K62S]; *Bank of England Mulls CBDC Models in Technology Engagement Forum*, LEDGER INSIGHTS (Oct. 21, 2021), https://www.ledgerinsights.com/bank-of-england-mulls-cbdc-models-in-technology-engagement-forum/ [https://perma.cc/TY94-DMEF]; Tom Farren, *Hong Kong Exploring CBDC as Part of Fintech Strategy*, COINTELEGRAPH (Oct. 4, 2021), https://cointelegraph.com/news/hong-kong-exploring-cbdc-as-part-of-fintech-strategy [https://perma.cc/8XGH-ZP3H].

19. *See G7 UK 2021*, GOV.UK, https://www.g7uk.org/ [https://perma.cc/9XWQ-3N75] (last visited Nov. 22, 2021) (including, in this instance Australia, India, South Korea, and South Africa).

20. *See Public Policy Principles*, *supra* note 6, at 4–5.

21. *Id.* at 7–8.

22. *Id.* (recognizing that CBDC "ecosystems" should "consider" how to ensure data privacy of consumers, and "be aligned to the progress being made towards international standards," while declining to create such standards).

23. *See* Waller, *supra* note 16, at 2–3.

24. While data privacy provisions exist in state laws, *see, e.g.*, California Consumer Privacy Act of 2018, CAL. CIV. CODE § 1798.100 (West 2018), and international data privacy standards, *e.g.*, CROSS BORDER PRIVACY RULES SYSTEM, http://cbprs.org/ [https://perma.cc/UPM6-VLGX] (last visited Nov. 22, 2021), a review of state law and international standards is beyond the scope of this Note. Additionally, a full survey of federal data privacy law is beyond the scope of a Note of this length.

concludes by proposing a mechanism to implement data privacy standards for CBDC use.

II. BACKGROUND

This section covers the basics of how blockchain technology works, and how a central bank digital currency might be implemented. It also discusses how a CBDC compares to other forms of digital currency and what benefits and risks they might contain.[25]

A. While a Central Bank Digital Currency May Be Built on a Blockchain, it is Distinct from Cryptocurrencies or Other Digital Ledger Tokens

While blockchain-based technologies are increasingly in the public eye, they remain an emerging area of technology and law. Accordingly, a non-technical primer of blockchain technologies, how the technology operates, what use cases exist, and what a CBDC is follows.

1. Blockchain Technologies Combine Existing Technologies into a Ledger-Based Tool for Storing and Distributing Information

Blockchain is not a new technology; rather, it is a novel combination of existing technologies to allow for the decentralized storage and distribution of information.[26] Blockchain combines concepts such as peer-to-peer networks (e.g., the Internet), cryptographic keys (used in many secure messaging systems), democratic consensus mechanisms, and digital signatures.[27] The resulting combination allows for a unique system of storing and distributing information: a decentralized ledger that shows up-to-date information on ownership of assets and how entities have interacted with each other over time (i.e., a ledger of transactions between blockchain participants).[28]

25. A full, technical discussion of how blockchains operate and all forms of digital currency is beyond the scope of a Note of this length. The following material provides a brief primer on essential principles.

26. CHRIS JAIKARAN, CONG. RSCH. SERV., R45116, BLOCKCHAIN: BACKGROUND AND POLICY ISSUES 1–2 (2018).

27. *Id.*; *see also* PRIMAVERA DE FILIPPI & AARON WRIGHT, BLOCKCHAIN AND THE LAW: THE RULE OF CODE 2–3 (2018).

28. *See* DE FILIPPI & WRIGHT, *supra* note 27, at 3 ("[B]lockchain technology supports decentralized, global value transfer systems that are both transnational and pseudonymous . . . . Governments across the globe are experimenting with blockchains to secure and manage critical public records, including vital information and titles or deeds to property.").

The Internet itself exists as a series of protocols that define how individuals can interact with it.[29] These protocols are a universal language that any computer or device must "speak" to access the Internet, and as such, they draw limits around what can or cannot be done by people interacting with the network.[30] Traditionally, governments have been able to implement layers of protocol that prohibit or enable certain actions, allowing for control over digital content and actions of citizens within their borders.[31] Blockchain technologies exist as another layer of protocol, analogous to another "application" layer.[32] Blockchain protocols allow individuals to interact with the defined protocol to access information on ownership and submit changes (generally being transactions between users) to that information.[33] Whether those changes are accepted relies on how the protocol of the blockchain is defined.[34]

Blockchains typically use peer-to-peer networks: a distributed network where participating computers connect with each other in a one-to-many relationship, rather than each computer connecting to a central server.[35] Each computer in the network communicates using the same blockchain protocol to validate, store, and distribute information.[36] By storing the complete transactional history of the blockchain on each participant computer, the network is resistant to change, and information is validated by consensus.[37] If a majority of participant computers validate a transaction stored in the ledger, that "block" of transactions is added to the "chain"—forming a comprehensive ledger of all previous transactions.[38] Accordingly, blockchains are largely autonomous, where changes to the blockchain are implemented through democratic consensus mechanisms rather than a central authority.[39]

---

29. *See* ALEXANDER R. GALLOWAY, PROTOCOL: HOW CONTROL EXISTS AFTER DECENTRALIZATION 38–39 (2004).

30. *See id.* at 46–47.

31. *See* DE FILIPPI & WRIGHT, *supra* note 27, at 50–51. *But cf.* Eric Hughes, *A Cypherpunk's Manifesto*, ACTIVISM.NET (Mar. 9, 1993), https://www.activism.net/cypherpunk/manifesto.html [https://perma.cc/3PD6-GWT4] (defining core values of the cypherpunk movement (a precursor movement to the development of cryptocurrencies), including advocating for privacy, freedom of information, the right to anonymity, and a lack of government monitoring and censorship).

32. *See* GALLOWAY, *supra* note 29, at 130.

33. *See* DE FILIPPI & WRIGHT, *supra* note 27, at 54–55.

34. *Id.*

35. *See id.* at 42–45.

36. *See id.*

37. *See id.* at 42 ("Underlying each blockchain-based network is a consensus mechanism that governs how information can be added to the shared repository. Consensus mechanisms make it possible for a distributed network of peers to record information to a blockchain, in an orderly manner, without the need to rely on any centralized operator . . . .").

38. *See id.* at 42–45.

39. *See* DE FILIPPI & WRIGHT, *supra* note 27, at 42–45, 147–48. A blockchain-based network and system can even lead to the creation of a decentralized autonomous organization (DAO), which is "a particular kind of decentralized organization that is neither run nor controlled by any person but entirely by code" and "generally consist[s] of a collection of smart contracts that do not have any 'owner.'" *Id.*

Blockchain users maintain a private key[40] (a lengthy alphanumeric code) known only to them.[41] As demonstrated in Figure 1, below, this private key is used as the input for a cryptographic algorithm to generate a public key, which is used to publicly sign transactions.[42] A public key is unique to a single private key, and users maintain it by keeping the private key confidential.[43] While a user may input a private key to access the blockchain, only the output of the cryptographic algorithm (the public key) is stored within the blockchain ledger.[44] Knowledge of the public key does not necessarily reveal any information about the user, although it may mean that an individual's transactions may be tracked if the public key is connected to the user.[45]

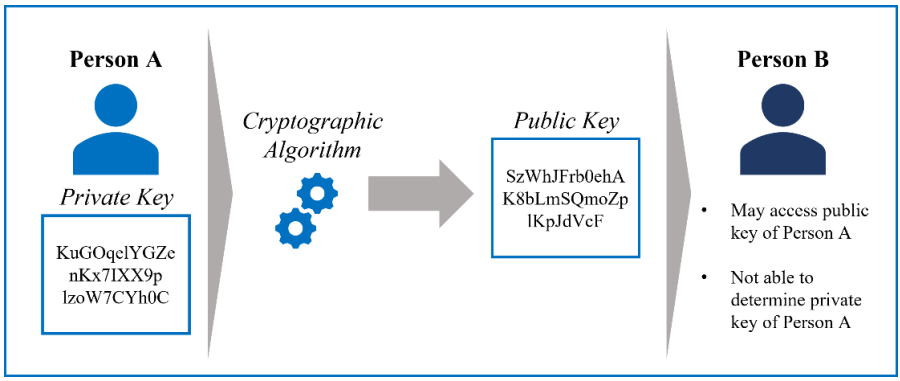

Figure 1: Private Keys Sign Transactions, Revealing a "Public Key" and Protecting User Anonymity

40. *See id.* at 38–39.

41. A private key is used in lieu of the more familiar process of signing into a secure website, wherein one enters a username and password, which is validated against a secure, central server. *See* Philip Bates, *How Do Websites Keep Your Passwords Secure?*, MUO (July 7, 2021), https://www.makeuseof.com/tag/websites-keep-passwords-secure/ [https://perma.cc/7P2S-UH5T]. The pitfalls of such a system are familiar: a compromised email address could be used to reset passwords associated with that email address, giving hackers access to multiple logins. *See id.* Alternatively, the central server could be hacked, compromising the username and password combinations of many users at once, unless otherwise encrypted. *See id.*

42. *See* JAIKARAN, *supra* note 26, at 1–2.

43. *Id.*

44. *Id.*

45. *Id.*

2. A Central Bank Digital Currency is Distinct from Cryptocurrencies and Stablecoins

While blockchains have many potential use cases,[46] one of the most common is in creating digital currencies, known as cryptocurrencies.[47] The fundamental information stored on a blockchain are digital assets or tokens, which have "money-like characteristics" and are used as a means of exchange for goods and services.[48] A user interacts with the blockchain by creating an account (often called a wallet) that has a private and public key.[49] The public key is used to create an address to which other users can send cryptocurrency tokens in some amount.[50] The blockchain stores the transactional data and proof of ownership in a series of coded "blocks," which maintains the informational integrity of the system: an anonymized, complete financial history of the transactions and interactions with the cryptocurrency blockchain.[51] Users participate in the blockchain by using their wallets to process transactions or by participating in "mining" of new blocks in the chain.[52] Mining blocks adds to the blockchain, allowing the transactional history of the network to continue to grow.[53] Mining rewards (e.g., tokens awarded for successfully "mining" a block) incentivize users to utilize the processing power of their computers[54] to help manage the decentralized blockchain.[55] Notably, in most blockchains, new tokens are only generated through mining and are not issued by a centralized body; anyone may participate, and anyone who participates may be rewarded for their successful participation.[56] A user typically acquires additional cryptocurrency by mining, by purchase on cryptocurrency exchanges, or by exchange (e.g., sale

46. *See, e.g.*, Jamie Berryhill et al., *Blockchains Unchained: Blockchain Technology and Its Use in the Public Sector* 13–15 (Org. for Econ. Coop. & Dev., Working Paper No. 28, 2018) (providing specific case studies for how blockchain technologies have been used in public sector applications (e.g., creation of a land registry to track ownership of land assets, inter-bank payments of international monetary or government securities transactions, or asset tracking for car ownership)).

47. *See* JAIKARAN, *supra* note 26, at 3, 5–6.

48. *See id.*

49. *See id.*

50. *See id.*; *see also CBDC vs Cryptocurrency*, *supra* note 2.

51. *See* DE FILIPPI & WRIGHT, *supra* note 27, at 42–45 (describing the term "blockchain"—a series of blocks "chained" together in series, creating an unchangeable, immutable ledger of past transactions using the blockchain-based system).

52. *See* Euny Hong, *How Does Bitcoin Mining Work?*, INVESTOPEDIA, https://www.investopedia.com/tech/how-does-bitcoin-mining-work/ [https://perma.cc/K52P-3NFX] (last visited Jan. 29, 2022).

53. *See id.*

54. Mining often utilizes specialized, networked equipment to "mine" on cryptocurrency networks. *See* Hong, *supra* note 52. A full discussion of the various hardware used for cryptocurrency blockchain engagement is not within the scope of this Note.

55. *See id.*

56. *See id.*

of goods or services in exchange for cryptocurrency).[57] The value of a cryptocurrency is dependent on the cost of production, the extent to which the community is involved in the blockchain, and general demand.[58] Accordingly, cryptocurrencies have been faulted for their extreme volatility.[59] Cryptocurrency blockchains also evolve over time as participants in the network participate in decentralized, democratic processes to vote on changes in the code.[60] If consensus is reached, the blockchain adopts the code-based amendment.[61]

Less common are stablecoins: cryptocurrencies with a value that is equivalent or "pegged" to a fiat currency (e.g., the U.S. dollar).[62] Stablecoins typically have a decentralized, peer-to-peer ledger and network, similar to cryptocurrencies.[63] The mechanism for creating additional stablecoins differs from cryptocurrencies: instead of creating new tokens by a user participation mining process, stablecoins are backed by fiat currency.[64] As users purchase stablecoins, the funds used to purchase the stablecoin are held as collateral to back the stablecoin, providing liquidity.[65] Often, cryptocurrency exchanges will use stablecoins as the necessary on- and off-ramps for users to exchange between cryptocurrencies and fiat currency; users exchange fiat currency for an equivalent amount of a stablecoin (e.g., U.S. dollar for the "USD coin" or USDC), and from there to cryptocurrency.[66] When users want to convert their cryptocurrency back to fiat currency, they are often forced to convert from cryptocurrency to stablecoin and finally back into fiat currency.[67] Apart from

57. *See id.*; *see also CBDC vs Cryptocurrency*, *supra* note 2. Additionally, participants in a blockchain-based system may acquire cryptocurrencies through a process known as "staking," wherein users lock up a certain amount of cryptocurrency in "validator" pools that earn interest over time. *See* Krisztian Sandor, *Crypto Staking 101: What Is Staking?*, COINDESK (Apr. 1, 2022, 11:37 AM), https://www.coindesk.com/learn/crypto-staking-101-what-is-staking/ [https://perma.cc/2HTA-264D].

58. *See* Hong, *supra* note 52; *see also CBDC vs Cryptocurrency*, *supra* note 2.

59. *See* Nicole Lapin, *Explaining Crypto's Volatility*, FORBES (Dec. 23, 2022, 6:00 AM), https://www.forbes.com/sites/nicolelapin/2021/12/23/explaining-cryptos-volatility [https://perma.cc/TJ9R-B87A].

60. *See What Are Blockchain Forks?*, CMC MARKETS, https://www.cmcmarkets.com/en/learn-cryptocurrencies/what-is-a-blockchain-fork [https://perma.cc/7ZKY-C5H8] (last accessed Sept. 29, 2022).

61. *See id.*

62. Fiat currency is a government-issued currency, specifically one not backed by a commodity (e.g., precious metals). *See* James Chen, *Fiat Money*, INVESTOPEDIA, https://www.investopedia.com/terms/f/fiatmoney.asp [https://perma.cc/VB4V-PZ8B] (last visited Mar. 5, 2022). Instead, the value backing the currency is the strength and stability of the government issuing the currency. *See id.*

63. *See* Adam Hayes, *Stablecoin*, INVESTOPEDIA, https://www.investopedia.com/terms/s/stablecoin.asp [https://perma.cc/A4R7-43YC] (last visited Apr. 9, 2022).

64. *See id.*

65. *See id.* (indicating that stablecoins may actually be either fiat-collateralized or crypto-collateralized, but in either form, some store of value is used as collateral and to provide liquidity as needed).

66. *See id.*

67. *See, e.g.*, *Withdrawals*, COINBASE, https://help.coinbase.com/en/commerce/getting-started/withdrawals [https://perma.cc/SLH3-W2XD] (last visited Sept. 29, 2022).

the mechanism of backing stablecoins with fiat currency, they are essentially indistinguishable from a cryptocurrency.[68]

A central bank digital currency (CBDC) differs from a cryptocurrency in four main ways: (1) the network model, (2) how the price of the token is determined, (3) the extent to which user information is stored on the blockchain, and (4) how changes to the blockchain are managed.[69] First, the network model in a CBDC is typically centrally managed, as opposed to a peer-to-peer, decentralized network.[70] The entire CBDC blockchain would be functionally under the control of the Federal Reserve, even if the development were outsourced to a third party.[71] The Federal Reserve would necessarily maintain an application programming interface (API) allowing CBDCs to be issued to commercial banks or directly to users.[72] Second, the price of the token is determined in the same way as fiat currency: through carefully managed monetary policy from the issuing authority.[73] In other words, a CBDC would be issued as "a digital liability of the Federal Reserve that is widely available to the general public."[74] A CBDC could be directly issued to other banks or private parties without mechanisms such as deposit insurance or backing by an underlying asset pool.[75] Third, use of a CBDC would necessarily require users to reveal personal information (e.g., the same information traditionally used to open a bank account: name, SSN, verification of identification, etc.).[76] A CBDC would also generate data about users' financial transactions and history, not unlike the financial data that is generated today.[77] However, this data would include the entire web of transactional data between users: showing how each CBDC came to be in each individual's wallet.[78] Such transaction history would theoretically be centralized with the blockchain manager (the Federal Reserve), even if minimally anonymized by using a public-private key encryption model

68. *See, e.g.*, Hayes, *supra* note 63 ("A stablecoin is a class of cryptocurrencies that attempt to offer price stability and are backed by a reserve asset. Stablecoins . . . offer the best of both worlds—the instant processing and security or privacy of payments of cryptocurrencies, and the volatility-free stable valuations of fiat currencies.").

69. *See CBDC vs Cryptocurrency*, *supra* note 2.

70. *See* Van Niekerk, *supra* note 4. Note that the introduction of various bills in Congress, including the Electronic Currency And Secure Hardware Act (ECASH Act) may—if passed into law—both (a) authorize the FRS to issue a CBDC and (b) prohibit the use of "a decentralized ledger (or indeed, any ledger of any type), which its proponents argue will help preserve user privacy." Nikhilesh De, *Lawmakers Keep Mentioning Privacy in CBDC Discussions*, COINDESK (Apr. 5, 2022, 5:16 PM), https://www.coindesk.com/policy/2022/04/05/lawmakers-keep-mentioning-privacy-in-cbdc-discussions/ [https://perma.cc/PW6S-LSR4]. It is unclear how a CBDC might be issued without any ledger system showing asset ownership. *Id.*

71. *See* Van Niekerk, *supra* note 4.

72. *See id.*

73. *See CBDC vs Cryptocurrency*, *supra* note 2.

74. MONEY AND PAYMENTS, *supra* note 13, at 13.

75. *See id.*

76. *See id.* at 13, 19.

77. *Id.* at 19.

78. *See* De, *supra* note 70.

similar to cryptocurrency.[79] Fourth, changes to the underlying protocol of a CBDC network would be determined and implemented by the issuing authority—the central bank—as opposed to a consensus-based user participation model.[80] Changes to the protocol would have a material (and potentially adverse)[81] impact on the user, as addressed below.[82]

B. A Central Bank Digital Currency Gives Powerful Monetary Policy Tools to the Government but Poses Inherent Privacy Risks

A central bank digital currency could substantially modernize our financial system.[83] Doing so could benefit consumers in the U.S. and maintain the strength of the U.S. dollar worldwide.[84] However, doing so without implementing effective safeguards could compromise consumer data.[85] Whether or not the benefits of a CBDC outweigh the risks remains to be seen; however, CBDC development is unlikely to begin until risks are adequately addressed.[86]

1. A Central Bank Digital Currency Provides Monetary Policy Tools to Ensure Equitable Access to Online Financial Payment Methods

The Board of Governors for the Federal Reserve System released a report in January 2022 detailing five benefits of a central bank digital currency.[87] First, a CBDC would "safely meet future needs and demands for payment services."[88] For example, as the economy increasingly goes "digital," a CBDC could be used as digital cash for online transactions.[89] A CBDC could lessen credit and liquidity risks to individual users by providing easy access to a digital "cash" form of money.[90] Instead of using credit or debit cards and accounts, consumers could directly pay for online transactions using a CBDC as digital cash (whereas the current system requires days or

79. *See* Van Niekerk, *supra* note 4.
80. *See id.*; *CBDC vs Cryptocurrency*, *supra* note 2.
81. *See, e.g.*, Tim Hakki, *Edward Snowden: CBDCs Are 'Cryptofascist Currencies' That Could 'Casually Annihilate' Savings*, DECRYPT (Oct. 10, 2021), https://decrypt.co/83124/edward-snowden-cbdcs-are-cryptofascist-currencies-that-could-casually-annihilate-savings [https://perma.cc/8XCG-WFU7] (highlighting concerns that "negative interest rates" could be used to encourage spending, which could be used as a tool to spur economic growth).
82. *See, e.g.*, MONEY AND PAYMENTS, *supra* note 13, at 17.
83. *See id.* at 13.
84. *See id.* at 15.
85. *See* Van Niekerk, *supra* note 4.
86. *See* MONEY AND PAYMENTS, *supra* note 13, at 19–20.
87. *See id.* at 14–16.
88. *Id.* at 14.
89. *Id.* at 15.
90. *See id.* at 14–15.

weeks to reconcile transactions).[91] Second, a CBDC could lead to "improvements to cross-border payments."[92] In fact, limited trials have shown that cross-border payments can be made using CBDCs in seconds, instead of the current "three to five days."[93] Not only would the time savings represent significant efficiency gains over the current system for cross-border payments, but using a CBDC would reduce the costs of such payments by up to 50%.[94] Third, a CBDC would "support the dollar's international role."[95] Recognizing that the dollar is widely used internationally, easy access to a CBDC could help ensure widespread use and adoption of the U.S. dollar (e.g., preventing decrease in dollar usage as other countries adopt easily accessible CBDC using their own currencies or CBDCs released by other nations).[96] Fourth, a CBDC could reduce barriers and lower transactional costs to "financial inclusion,"[97] benefitting low-income and unbanked households.[98] Fifth, a CBDC would "extend public access to safe central bank money," especially in an increasingly digital world.[99] Use of a CBDC would provide the online equivalent to using cash online, rather than relying on traditional payment systems which carry credit and liquidity risks.[100]

Use of a CBDC places monetary tools into the hands of the Federal Reserve System to accomplish the benefits described above.[101] Choices in the design and implementation of a CBDC would affect how users perceive and use a CBDC system.[102] For example, the amount of interest a CBDC would accrue could be changed at will to encourage spending or saving as a tool against inflation.[103] Protocols could also facilitate the rapid payment of taxes, tax refunds, delivery of wages, and access to credit.[104] Possibly some of these additional features could drive adoption of a CBDC; some users who might

91. *See id.*

92. MONEY AND PAYMENTS, *supra* note 13, at 15.

93. Alun John, *Central Bank Digital Currencies Can Slash Cross Border Payment Time*, REUTERS (Sept. 28, 2021, 3:07 AM), https://www.reuters.com/business/central-bank-digital-currencies-can-slash-cross-border-payment-time-bis-2021-09-28/ [https://perma.cc/7NPY-LHDT].

94. *See id.*

95. MONEY AND PAYMENTS, *supra* note 13, at 15.

96. *See id.*

97. *Id.* at 16.

98. *Id.* (stating that further study is necessary to assess the potential for CBDC to help "underserved and lower income households"). *Contra* Waller, *supra* note 16, at 2–3 (suggesting that less than 1% of American households are both unbanked *and* potentially interested in a CBDC account issued by the Federal Reserve System).

99. MONEY AND PAYMENTS, *supra* note 13, at 16.

100. *See id.* (describing how cash use in the United States has decreased from 40% of transactions in 2012 to 19% of transactions in 2020, a trend that is likely to continue).

101. *See id.* at 16–17.

102. *See id.* at 17.

103. *See, e.g.*, *id.* (suggesting that a "non-interest-bearing CBDC" could make CBDC use "less attractive as a substitute for commercial bank money" and therefore limit changes to the traditional financial-sector); *see also* Hakki, *supra* note 81.

104. *See* MONEY AND PAYMENTS, *supra* note 13, at 16.

not see the utility in "digital cash" may nevertheless use a CBDC if it provides an easier way to handle taxes or access credit.[105]

2. Data Aggregation Creates Significant Risk for Consumer Data Privacy

The Federal Reserve System paper on CBDCs flags "complex policy issues and risks" that could benefit from additional scholarship and analysis.[106] CBDC usage could lead to widespread "changes to financial-sector market structure[s]."[107] Banks traditionally rely on central bank deposits to fund loans to consumers; a CBDC would provide direct competition with commercial bank money and could result in "increased bank funding expenses . . . and reduce credit availability or raise credit costs for households and businesses" as the aggregate value of central bank deposits in commercial banks decreases.[108] Direct consumer access to a CBDC could make "runs on financial firms more likely or severe," undercutting safeguards currently in place to prevent bank runs.[109] Over time, to the extent that CBDCs provide simplified access to credit options, use of commercial banks could decline precipitously, especially given the increasing digitization of commerce.[110]

An important area of risk for the Federal Reserve is ensuring "privacy and data protection and the prevention of financial crimes."[111] There is a balancing act between the necessity of preventing financial crimes and the necessity of data privacy and protection.[112] Perfect financial information would all but negate the possibility for financial crimes, whereas complete anonymity would afford protection of consumer data but provides ample

105. *See* Waller, *supra* note 16, at 2–3.

106. MONEY AND PAYMENTS, *supra* note 13, at 17. A full discussion of all of the complex policy issues and risks contained in the FRS paper is beyond the scope of this Note: indeed, additional scholarship is needed to continue to address the potential risks of a CBDC system.

107. *Id.*

108. *Id.* (suggesting also that the increase in cryptocurrency and stablecoin use poses similar risks to commercial banks). *Contra* Hughes, *supra* note 31 (defining the radical transformation of traditional systems, which forms the core of the cypherpunk movement—leading to the initial development of cryptocurrencies: such a transformation to the financial-sector market structure is in-line with the earliest goals of the cryptocurrency movement).

109. MONEY AND PAYMENTS, *supra* note 13, at 17.

110. *See, e.g.*, *id.*

111. *Id.* at 19.

112. *See* MONEY AND PAYMENTS, *supra* note 13, at 19.

ground for the growth and proliferation of underworld financial schemes.[113] Some level of collection of consumer data is essential with a CBDC to support anti-money laundering (AML) policy goals and would likely involve similar data to what is now collected from consumers in opening a bank account.[114]

The Federal Reserve System waves aside such concerns, stating that an intermediary system would be used to issue CBDCs, and those intermediaries (i.e., commercial banks) would utilize "existing tools" to collect and protect consumer data.[115] This argument ignores a fundamental conflict of interest: CBDC funds operate in direct competition with commercial bank funds, offering limited incentive for commercial banks to offer CBDC accounts to users.[116] For this very reason, many countries are likely to adopt a direct-to-consumer CBDC issuance system.[117] Such a system necessarily requires the "digitization and centralization of identity" to verify user information and limit the possible commission of financial crimes.[118] This places personally identifiable information in the hands of the Federal Reserve System and then connects that information explicitly to the spending habits and practices of individuals.[119]

Even in an intermediated system where personally identifiable information is not maintained by the Federal Reserve, the data privacy risks posed by a CBDC are expansive. The Federal Reserve would have access to an unprecedented aggregation of consumer financial data, including a ledger showing the complete and accurate ownership of all assets by account, as well as a list of every transaction from account to account.[120] This would allow the tracing of a single CBDC dollar from issuance to the current account holder.[121] Imagine that the government knew not only how much money was in your wallet, but the serial numbers of every dollar bill in your wallet and how it came to be there.[122] This is such a radical shift from the current baseline

113. *See id.* Fears that criminals might want to use a CBDC are overstated. *See* Tom Sadon, *5 Reasons Why Criminals & Terrorists Turn to Cryptocurrencies*, COGNYTE (Nov. 2, 2021), https://www.cognyte.com/blog/5-reasons-why-criminals-are-turning-to-cryptocurrencies/ [https://perma.cc/PXY3-ABH3] (stating that criminals may use cryptocurrencies because they offer some privacy, are not centrally managed, can process transactions quickly, and are borderless). While criminals occasionally use cryptocurrencies, their reasons for doing so are, in effect, the list of differences between a cryptocurrency and a CBDC. *See id.* A CBDC has no such promise of anonymity or privacy and is centrally managed by the U.S. government—which tends to support a preliminary hypothesis that a CBDC would not be attractive to the criminal underworld. *See, e.g.*, *id.*

114. *See* MONEY AND PAYMENTS, *supra* note 13, at 17–18.

115. *See id.* at 13–14, 17.

116. *See id.* at 17, 19.

117. *See Central Bank Digital Currency Tracker*, *supra* note 1.

118. Van Niekerk, *supra* note 4.

119. *See, e.g.*, *id.* (detailing the connection from digitization and centralization of identity to use of CBDC systems as a method for digital signatures, access to government services, and linking payments to individual identity).

120. *See id.*

121. *See CBDC vs Cryptocurrency*, *supra* note 2.

122. *See id.*; Van Niekerk, *supra* note 4.

that it is not considered by current data privacy law.[123] Granted, in an intermediary system, such data may be anonymized,[124] but the personal nature of spending habits is a factor in some transactions that remain in cash today.[125] Even when anonymized, use of a CBDC would place the entire web of financial transaction data in the hands of the federal government, and "with great power comes great responsibility"—in this case, the need to create robust federal data privacy protections.[126]

III. FEDERAL FINANCIAL DATA PRIVACY: THE GRAMM LEACH-BLILEY ACT

United States data privacy law is a multijurisdictional patchwork of state and federal laws.[127] The most significant federal law establishing data privacy standards for financial institutions is the Gramm-Leach-Bliley Act ("GLBA").[128] While some state laws may exceed the data privacy standards in the GLBA,[129] these state laws cannot be enforced against the federal government.[130] Some state laws, like the California Consumer Privacy Act, may provide a helpful model for expanding federal data privacy protections to consumers.[131] However, state laws are less relevant to a discussion of the issuance of central bank digital currencies by the Federal Reserve—a federal agency and accordingly, an in-depth discussion of state data privacy law is out of scope for this Note.[132] This section proceeds with an analysis of the GLBA: its history and legislative purpose, relevant data privacy provisions, and the applicability of the GLBA to federal agencies as financial institutions.The Gramm-Leach-Bliley Act was signed into law in 1999 in an effort to "enhance competition in the financial services industry by providing a prudential framework for the affiliation of banks, securities firms, insurance

123. *See* Soubouti, *supra* note 15, at 534–35.

124. *See* MONEY AND PAYMENTS, supra note 13, at 19.

125. *See id.* at 16; Waller, *supra* note 16, at 4.

126. *See* Aaron Gleason, *Steve Ditko's Great Gift to the World: 'With Great Power Comes Great Responsibility'*, FEDERALIST (July 9, 2018), https://thefederalist.com/2018/07/09/steve-ditkos-great-gift-world-great-power-comes-great-responsibility/ [https://perma.cc/KQ8Z-5LXB] (describing the origins of the phrase as likely dating to the allegory of the Sword of Damocles—perhaps another apt metaphor for the data privacy concerns posed by a CBDC); *see also* Van Niekerk, *supra* note 4.

127. *See* Soubouti, *supra* note 15, at 527–28.

128. *See id.* at 528–29.

129. *See id.* at 531.

130. McCulloch v. Maryland, 17 U.S. 316, 426 (1819) ("This great principle is, that the constitution and the laws made in pursuance thereof are supreme; that they control the constitution and laws of the respective states, and cannot be controlled by them.").

131. *See* Meredith E. Bock, Note, *Biometrics and Banking: Assessing the Adequacy of the Gramm-Leach-Bliley Act*, 24 N.C. BANKING INST. 309, 321–22 (2020); California Consumer Privacy Act of 2018, CAL. CIV. CODE § 1798.140(b) (West 2018).

132. *See* Bock, *supra* note 131, at 321–22.

companies, and other financial service providers"[133] The GLBA applies to "financial institutions," creating an affirmative duty to "respect the privacy of its customers" and to protect customer "nonpublic personal information."[134] "Nonpublic personal information" is defined as "personally identifiable financial information (i) provided by a consumer to a financial institution; (ii) resulting from any transaction with the consumer or any service performed for the consumer; or (iii) otherwise obtained by the financial institution."[135]

Exceptions to "nonpublic personal information" exist for information that is publicly accessible.[136] In other words, a consumer may expect that financial institutions (such as a bank) will safeguard any personal information she explicitly provides (including, e.g., name, date of birth, SSN, address, income information)[137] as well as information related to transactions with the bank.[138] A consumer using a credit card provided by a commercial bank, therefore, should not expect that any transactions using the credit card are private.[139] However, a consumer using cash withdrawn from a bank ATM may expect that any transactions using that cash are private; the bank is only aware of the fact that a certain amount of cash was withdrawn at an ATM by that user, not what happens to the cash after the fact.[140] Financial institutions must provide privacy and opt-out notices to inform customers of data privacy policies and provide a mechanism for individuals to opt-out of a financial institution sharing information with "nonaffiliated third parties."[141] Financial institutions must also maintain customer data safely and securely.[142]

The GLBA has been held to apply to federal institutions such as "credit reporting agencies."[143] Indeed, the text of the GLBA states that it applies to each "agency or authority" that is a "financial institution."[144] A "financial institution" includes any institution engaged in "financial activities," excluding institutions that do not "sell or transfer nonpublic personal

133. Individual Reference Servs. Grp., Inc. v. FTC, 145 F. Supp. 2d 6, 17 (D.D.C. 2001), *aff'd sub nom.* Trans Union LLC v. FTC, 295 F.3d 42 (D.C. Cir. 2002) (quoting H.R. REP. NO. 106-434, at 245 (1999) (Conf. Rep.)).

134. 15 U.S.C. § 6801(a).

135. 15 U.S.C. § 6809(4)(A).

136. 15 U.S.C. § 6809(4)(B).

137. *See* Bock, *supra* note 131, at 315.

138. *See* 15 U.S.C. § 6801(a).

139. *See, e.g.*, Bock, *supra* note 131, at 315.

140. *See, e.g.*, Brad Berens, *Why Using Cash Won't Protect Your Privacy*, CTR. DIGIT. FUTURE (Jan. 4, 2018), https://www.digitalcenter.org/columns/cash-and-anonymity/ [https://perma.cc/R3UH-R7XJ]. Such an analogy breaks down when the cash is replaced with a digital token that is tracked. If either a commercial bank or the Federal Reserve System is aware of every single issued CBDC "dollar"—where it is, how it got there, and who currently owns it—then either institution has access to data that was not considered under the GLBA or other federal data privacy laws. *See* 15 U.S.C. § 6809(4)(A).

141. *See* Bock, *supra* note 131, at 315–16.

142. *See* Bock, *supra* note 131, at 315–16.

143. *Individual Reference*, 145 F. Supp. 2d at 17 (quoting H.R. REP. NO. 106-434, at 245).

144. 15 U.S.C. § 6801.

information to a nonaffiliated third party."[145] The FRS does not currently collect consumer data. In fact, the Federal Reserve Act "does not authorize direct Federal Reserve Accounts for individuals, and such accounts would represent a significant expansion of the Federal Reserve's role in the financial system and the economy."[146] The issue of whether the GLBA applies to the FRS is therefore currently moot.[147] However, if individuals were issued CBDC funds directly from the FRS, the FRS would undoubtedly fall under and be required to follow the requirements of the GLBA.[148]

The GLBA does not require financial institutions to safeguard consumer data that is not protected by the Act.[149] This includes information gathered on websites from visitors or non-customers, including "behavioral biometric data."[150] Behavioral biometric data includes keystrokes and navigation of a webpage to verify a user's identity; such data can create a unique user profile to identify users who do not provide data otherwise covered by the GLBA.[151] This kind of data is currently used in fraud detection by financial institutions to highlight anomalous customer behavior.[152] CBDC data could provide a similar "user profile" constructed of all of a user's transactions using digital cash.[153] Such data would contain an interwoven mixture of protected and unprotected data.[154] To the extent that data is not currently protected by the GLBA, financial institutions may have little incentive to safeguard user data. Accordingly, as discussed below, the scope of the GLBA should be amended to include the data types that would be collected in use of a CBDC.

IV. PROPOSED DATA PRIVACY STANDARDS FOR CENTRAL BANK DIGITAL CURRENCIES

A. The Federal Reserve System Has Not Addressed Data Privacy Concerns Inherent in CBDCs

Use of a CBDC would necessarily involve the widespread collection and use of consumer data.[155] As discussed previously, consumers would not only furnish the types of data used in setting up a bank account to initially set up a wallet for CBDC use, but would also necessarily consent to the collection

145. 15 U.S.C. § 6809(3).
146. MONEY AND PAYMENTS, *supra* note 13, at 19.
147. *See id.* at 13–14.
148. 15 U.S.C. § 6809(3).
149. *See, e.g.*, 15 U.S.C. § 6809(4).
150. *See* Soubouti, *supra* note 15, at 534-35; Bock, *supra* note 131, at 313.
151. *See* Bock, *supra* note 131, at 313.
152. *See id.*
153. *Id.* at 313; *see also* Van Niekerk, *supra* note 4; *CBDC vs Cryptocurrency*, supra note 2.
154. *See, e.g.*, Van Niekerk, *supra* note 4.
155. *See id.*

of all transaction data.[156] Such data is not siloed by the customer; the CBDC ledger would show the entire financial web of transactions from customer to customer—thus providing a perfect, up-to-date ledger of CBDC ownership and history for all customers.[157] This three-dimensional data is not contemplated within the GLBA's definition of "nonpublic personal information."[158] To adequately safeguard such data, Congress should amend the GLBA to more explicitly define protected data to include that which would be collected in the routine course of CBDC use.[159] To the extent that the FRS engages with this data, the GLBA also should be amended to explicitly incorporate the FRS as a financial institution, and the FRS should in turn work to develop the institutional competence and tools necessary to adequately safeguard consumer data.[160]

The FRS, for its part, denies that it would collect data in issuing a CBDC.[161] They instead point to an intermediated model, which would allow the FRS to issue CBDC funds to commercial banks, who in turn would offer "accounts or digital wallets" to users to "facilitate the management of CBDC holdings and payments."[162] However, this argument misses the mark for two reasons. First, a CBDC would necessarily be built on a centralized blockchain managed by FRS.[163] Although commercial bank accounts could facilitate the *management* of CBDC holdings and payments, the underlying financial data—who owns what at any given moment—would be stored *at and by* the FRS.[164] Commercial banks, bound as they are by anti-money-laundering and data privacy laws, would still be required to collect the same information to open a CBDC account as they would for any other bank account: the status quo.[165] Yet, the FRS would maintain control over the bulk of financial data inherent in the CBDC system: a dramatic departure from the status quo unaddressed by the FRS.[166]

---

156. *See* Van Niekerk, *supra* note 4; Soubouti, *supra* note 15, at 534–35.

157. *See* Van Niekerk, *supra* note 4 (stating that a CBDC could "[b]e tracked across every movement, where the account that is credited appends that information to the digital dollar, in perpetuity" and "[b]e stopped, returned to the source, returned to the previous account, or even destroyed at any moment.").

158. *See* Soubouti, *supra* note 15, at 534–35. The data is three-dimensional in the sense that for a single transaction, the data could show the relationship between the FRS and each party to the transaction, the relationship between parties to the transaction itself, and the relationships between each party to the transaction and all third parties with whom parties have transacted leading up to the transaction being examined. *Id.*

159. *See* Bock, *supra* note 131, at 326. The FRS has indicated that they will not implement a CBDC without direct authorization and support from Congress. *See* MONEY AND PAYMENTS, *supra* note 13, at 3 ("The Federal Reserve does not intend to proceed with issuance of a CBDC without clear support from the executive branch and from Congress, ideally in the form of a specific authorizing law."). Accordingly, any such authorization should include, as part and parcel, adequate data privacy standards in the form of a modification to the GLBA.

160. *See, e.g.*, Van Niekerk, *supra* note 4.

161. *See* MONEY AND PAYMENTS, supra note 13, at 13–14.

162. *Id.*

163. *See* Van Niekerk, *supra* note 4.

164. *See id.*

165. *See* MONEY AND PAYMENTS, *supra* note 13, at 19.

166. *See* Van Niekerk, *supra* note 4.; MONEY AND PAYMENTS, *supra* note 13, at 19.

Second, this argument ignores the fact that CBDC funds would operate in direct competition with commercial bank funds.[167] Commercial banks have no financial incentive to offer access (by extending credit options or otherwise) to a digital cash system that would reduce their profitability by funneling activity away from their own online transaction services.[168] To solve this issue, either some additional incentive would need to be provided to commercial banks to provide access to CBDC accounts for users, or the federal government (likely the FRS as owner of the CBDC system, network, and protocol) would need to step in to provide public access to consumers interested in opening a CBDC account.[169] Assuming that CBDC is legal tender, all businesses would have to accept CBDC funds and would therefore need a CBDC account, requiring the FRS to quickly develop the capability to handle millions of accounts.[170]

B. Solutions to Protect Consumer Data Privacy Include Commercial Bank Incentives, FRS Reform, and Legislation to Expand the Gramm-Leach-Bliley Act

To ensure that consumer data privacy is adequately safeguarded, there are three potential solutions.[171] First, a CBDC should be designed to incentivize commercial banks to make available CBDC accounts.[172] In an intermediated system, such as that proposed by the FRS in their *Money and Payments* paper, bank provision of CBDC accounts would not represent a significant expansion in data collected by such banks; commercial banks already collect this data routinely.[173] However, as discussed above, banks have little incentive to provide accounts that act in direct competition with

167. *See* MONEY AND PAYMENTS, *supra* note 13, at 17.

168. *See id.*; *see also* Van Niekerk, *supra* note 4.

169. *See* MONEY AND PAYMENTS, *supra* note 13, at 17 (also stating the risk of increased use of stablecoins in lieu of CBDC if such accounts are not generally available).

170. *See* James B. Thayer, *Legal Tender*, 1 HARV. L. REV. 73, 73 (1887) (discussing the history of legal tender at the foundation of our country, which strongly mirrors the debate over whether the FRS may issue a CBDC); *see also* Jess Cheng & Joseph Torregrossa, *A Lawyer's Perspective on U.S. System Payment Evolution and Money in the Digital Age*, BD. OF GOVERNORS OF THE FED. RSRV. SYS. (Feb. 4, 2022), https://www.federalreserve.gov/econres/notes/feds-notes/a-lawyers-perspective-on-us-payment-system-evolution-and-money-in-the-digital-age-20220204.htm [https://perma.cc/EJL6-TCGD] (detailing the differences between a Federal Reserve note and a bank deposit, including the ability for commercial banks to "affect the total stock of money through lending activities that credit the accounts of borrowers" and "expose[] their balance sheet to risk.").

171. The following solutions are mutually exclusive but not collectively exhaustive. All three should be pursued in order to mitigate the data privacy risks inherent in a CBDC. However, it may be the case that additional solutions recommend themselves as the issues surrounding a CBDC in the United States are further studied through additional research and scholarship.

172. *E.g.*, MONEY AND PAYMENTS, *supra* note 13, at 17.

173. *See* Bock, *supra* note 131, at 315.

commercial bank funds.[174] Such incentives could take many forms: for example, there could be significant demand for user accounts, which could provide an incentive for commercial banks to offer CBDC accounts as a means for capturing greater market share.[175] Alternatively, Congress could provide monetary incentive for banks to offer user accounts, or a U.S. CBDC could be designed with the goal of ensuring "little to no disruption to the banking sector."[176]

Second, the FRS should begin to develop the institutional competence to safeguard consumer data. Such data could be limited to the underlying financial data inherent in a CBDC (i.e., the entire web of transactions).[177] However, if commercial banks are unwilling to offer CBDC accounts, this data could include the same types of data that are currently collected by banks and other financial institutions *in addition to* the underlying financial data inherent in a CBDC.[178] Beyond the protection of data, absent an intermediated system in which commercial banks offer user accounts, the FRS would need to develop infrastructure to support customers, which would likely include a variety of support services such as customer service centers, technical support, and other auxiliary support mechanisms.[179]

Third, the GLBA should be expanded to explicitly cover both the types of data that would be collected with a CBDC and the federal institutions involved in issuing and managing the data underpinning the CBDC system.[180] Whether or not an intermediated system is used to issue CBDC funds, the Federal Reserve would, as discussed above, maintain financial data showing every transaction on the CBDC system and could theoretically combine that data with personally identifiable information provided by consumers in opening a CBDC wallet or account.[181] These three-dimensional financial data types are not considered in the GLBA or other federal data privacy laws—a

174. *See* MONEY AND PAYMENTS, *supra* note 13, at 17; *see also* Van Niekerk, *supra* note 4.

175. *See* Jess Cheng et al., *Preconditions for a General-Purpose Central Bank Digital Currency*, BD. GOVERNORS FED. RSRV. SYS. (Feb. 24, 2021), https://www.federalreserve.gov/econres/notes/feds-notes/preconditions-for-a-general-purpose-central-bank-digital-currency-20210224.htm [https://perma.cc/2D5U-U3Q3].

176. *Id.*

177. *See* Van Niekerk, *supra* note 4.

178. *See id.*; *CBDC vs Cryptocurrency*, *supra* note 2.

179. Little scholarship addresses the point of developing institutional competence to handle such massive amounts of financial data. However, these competencies likely exist across government (e.g., financial data managed and stored by the IRS, or customer support call centers at GSA) from which the FRS could extract best practices in data management and customer support. Further research should be done to assess the technical and logistical requirements necessary to implement a CBDC, with care taken to identify the competencies that can reasonably be leveraged from across government.

180. *See, e.g.*, Bock, *supra* note 131, at 326.

181. *See* Van Niekerk, *supra* note 4; MONEY AND PAYMENTS, *supra* note 13, at 19.

gap that must be addressed prior to the development and implementation of a CBDC system.[182]

V. CONCLUSION

A central bank digital currency represents a substantial opportunity to "fundamentally change the structure of the U.S. financial system" to make it more equitable, accessible, and responsive to a modern and increasingly digital world.[183] A CBDC would bring the U.S. dollar into the modern world and ensure the longevity of the dollar's international role.[184] However, a CBDC brings inherent data privacy risks that are not considered under current federal data privacy laws; consumer identity would be linked to every single transaction made, offering a complete big data picture of the entire digital financial system.[185] An expansion of the GLBA to explicitly include the types of data that would be collected by a CBDC system, including underlying financial information that would comprise the CBDC blockchain, is necessary to ensure adequate safeguards for consumer data. As the Federal Reserve System continues to seek feedback on CBDC, more research is needed to further examine potential data privacy risks.[186]

182. *See, e.g.*, Soubouti, *supra* note 15, at 534–35 (discussing types of data that are not considered within the framework of the GLBA).

183. MONEY AND PAYMENTS, *supra* note 13, at 17.

184. *See id.*

185. *See* Van Niekerk, *supra* note 4.

186. *See* MONEY AND PAYMENTS, *supra* note 13, at 21 (indicating that "[t]he Federal Reserve will only take further steps toward developing a CBDC if research points to benefits for households, businesses, and the economy overall that exceed the downside risks, and indicates that CBDC is superior to alternative methods."). It remains to be seen whether the United States will officially determine whether to pursue development of a CBDC, and such an effort would likely take years to implement.

Do Androids Defame with Actual Malice? Libel in the World of Automated Journalism

Dallin Albright*

TABLE OF CONTENTS

* J.D., May 2023, The George Washington University Law School; B.S., Geology, Brigham Young University. Thank you to my colleagues on the Federal Communications Law Journal for their hard work and support in the publication process. Special thanks to Michael Beder, Journal Adjunct, for his guidance and feedback in writing this Note.

I. INTRODUCTION

Automation has been a disruptive influence for many professions, and now even journalists are facing the effects. Automated journalism is the use of artificial intelligence (AI), or algorithmic computer programs, to produce news articles.[1] It has been used effectively by news outlets such as *The Washington Post*, *The Associated Press*, and *The New York Times* in sports scores, financial news, and reporting the weather.[2] In September 2020, *The Guardian* published a long-form article produced by OpenAI's GPT-3 language generator, demonstrating the potential of automated journalism.[3] Microsoft announced in 2020 that it would not renew contracts with roughly fifty of its news production contractors and that it planned to use AI to replace them.[4] In the next several years, AI is expected to transform the news industry, presenting novel legal challenges to those practicing communications law.[5]

Automated journalism creates a unique risk to news publishers with respect to the possible production of defamatory or libelous statements.[6] Courts in the past have created standards dependent on an author-defendant's malice or their understanding that a defamatory statement is false or hurtful.[7] However, traditional methods cannot show that an algorithm possessed malice or that a machine produced a statement knowing it was false or hurtful.[8] And yet, AI-generated defamation is still harmful to the individuals about whom it is written and to the general public consuming the false information.[9] Some argue that statements produced by an algorithm are owed

---

1. Andreas Graefe, *Guide to Automated Journalism,* TOW CTR. FOR DIGIT. JOURNALISM (Jan. 7, 2016), https://www.cjr.org/tow_center_reports/guide_to_automated_journalism.php [https://perma.cc/Z4PX-Q24H]; *Here Come the Writing Robots: How Is Automated Journalism Impacting the Media?*, TECHSLANG (Nov. 12, 2020), https://www.techslang.com/how-is-automated-journalism-impacting-the-media/ [https://perma.cc/ES9D-6D5B].

2. Corinna Underwood, *Automated Journalism – AI Applications at New York Times, Reuters, and Other Media Giants*, EMERJ ARTIFICIAL INTEL. RSCH. (Nov. 17, 2019), https://emerj.com/ai-sector-overviews/automated-journalism-applications/ [https://perma.cc/QAE5-JXBB].

3. GPT-3, *A Robot Wrote This Entire Article. Are You Scared Yet, Human?*, GUARDIAN (Sept. 8, 2020, 4:45 AM), https://www.theguardian.com/commentisfree/2020/sep/08/robot-wrote-this-article-gpt-3 [https://perma.cc/ZA9F-WT4X].

4. Geoff Baker, *Microsoft Is Cutting Dozens of MSN News Production Workers and Replacing Them with Artificial Intelligence,* SEATTLE TIMES (May 29, 2020, 8:43 PM), https://www.seattletimes.com/business/local-business/microsoft-is-cutting-dozens-of-msn-news-production-workers-and-replacing-them-with-artificial-intelligence/ [https://perma.cc/8Y8Z-HA4Z].

5. Ron Schmelzer, *AI Making Waves in News and Journalism,* FORBES (Aug. 23, 2019, 10:48 AM), https://www.forbes.com/sites/cognitiveworld/2019/08/23/ai-making-waves-in-news-and-journalism/ [https://perma.cc/8Y8Z-HA4Z].

6. Seth C. Lewis et al., *Libel by Algorithm? Automated Journalism and the Threat of Liability,* 96 JOURNALISM & MASS COMM. Q. 60, 61 (2019).

7. *See* New York Times Co. v. Sullivan, 376 U.S. 254, 279-80 (1964).

8. Lewis et al., *supra* note 6, at 68.

9. *See* Pascal Podvin, *The Social Impact of Bad Bots and What to Do About Them,* FORBES: TECH. COUNCIL (Dec. 4, 2020, 9:00 AM), https://www.forbes.com/sites/forbestechcouncil/2020/12/04/the-social-impact-of-bad-bots-and-what-to-do-about-them/ [https://perma.cc/BW4E-SFEB].

the same protections afforded to the statements made by living individuals.[10] Others believe that as non-human actors, algorithms do not warrant the same level of protection as human speakers.[11]

This Note argues that the actual malice standard for defamation should not apply to statements produced by AI, even when the statements discuss public officials or public figures. Rather, defamation claims for AI-generated statements should be evaluated under the more appropriate negligence standard, which is usually applied to statements about private individuals. Under the negligence standard, defendants would have a reasonable duty of care to follow journalistic practices and attempt to ascertain the truthfulness of statements generated by AI. This is more appropriate than the actual malice standard, which requires only that a defendant did not have serious doubts about a statement's truthfulness and was not recklessly indifferent in publishing them.

This Note will first review the nature and development of algorithmic speech before analyzing how the negligence standard could be applied to cases involving AI. The Background section will review how algorithms create statements through mechanical patterns with various degrees of human input, and how this process can sometimes lead to unpredictable results. This section will also review the elements of libel law, demonstrating the unique protection given to defendants who make statements about public officials and public figures on account of a constitutional concern for freedom of speech. The Analysis section will then examine the reasoning behind imposing a stricter duty upon defendants that use AI on account of its unique power to spread disinformation if left unchecked. Then, this Note will address concerns that free speech advocates may have against removing the actual malice requirement by analyzing the difference between algorithmic speakers and human speakers. AI poses a unique challenge to legal and journalistic institutions, and only by adapting quickly can courts keep up with rapidly developing technology.

II. BACKGROUND

To understand the reasons for removing the actual malice requirement for libel when speech is produced by AI, it is necessary to understand the basic nature of artificial intelligence and the legal framework surrounding defamation. Autonomous journalism currently requires significant human input, but as the technology becomes more sophisticated, it will require less and less independent human judgment to create and share statements.[12] This can lead to false, inappropriate, or misleading statements being shared with the public if not properly reviewed or controlled.[13] The elements of libel

10. *See* Toni M. Massaro & Helen Norton, *Siri-ously? Free Speech Rights and Artificial Intelligence*, 110 NW. U. L. REV. 1169, 1178-79 (2016) (quoting JOEL FEINBERG, FREEDOM AND FULFILLMENT: PHILOSOPHICAL ESSAYS 52 (1992)).

11. *See, e.g.*, Oren Bracha & Frank Pasquale, *Federal Search Commission? Access, Fairness, and Accountability in the Law of Search*, 93 CORNELL L. REV. 1149 (2008).

12. *See* Graefe, *supra* note 1.

13. *See* Podvin, *supra* note 9.

against public figures require that, in addition to the statement being untrue, a defamatory statement is shared with actual malice or reckless disregard for the truth.[14] This could create a difficult barrier for those damaged by autonomously-generated libel to overcome because algorithms cannot be shown to possess actual malice or reckless disregard for the truth in the same way human authors can possess.

A. Algorithmic Speech

Statements produced by AI are commonly called "algorithmic speech," and can be classified in several broad categories based on the level of user input required to produce statements.[15] This Note will adopt the categories of Curated Production, Semi-Autonomous Production, and Fully Autonomous Production.[16] Before addressing legal challenges presented by speech produced by AI, it is essential for this Note to define and describe these categories of speech.

1. Curated Production

Curated production is a form of algorithmic speech where computer programs are fed data internally by users to produce text.[17] This level of AI possesses less freedom to generate unexpected statements and the greatest amount of user control.[18] Programs like these are fed information to produce text that is formulaic and predictable.[19]

Most current autonomously-generated news stories would be categorized as Curated Production.[20] News companies feed a program data from sports matches, weather forecasts, or the financial markets, and the program produces simple stories that resemble those written by a human.[21] Since these news stories are mostly "by-the-numbers" with little to no commentary or analysis, they are ideal for autonomous journalism, and many news publishers have adopted the technology specifically to cover these fields.[22]

14. *See* New York Times Co. v. Sullivan, 376 U.S. 254, 279-80 (1964).

15. *See* Alan M. Sears, *Algorithmic Speech and Freedom of Expression*, 53 VAND. J. TRANSNAT'L L. 1327, 1333-34 (2020).

16. *Id.*

17. *Id.* at 1333.

18. *See id.* at 1333-34.

19. Stephen Beckett, *Robo-Journalism: How a Computer Describes a Sports Match,* BBC: CLICK TV (Sept. 12, 2015), https://www.bbc.com/news/technology-34204052 [https://perma.cc/3Q2B-DJJA].

20. *See* Sears, *supra* note 15, at 1333.

21. Graefe, *supra* note 1.

22. *Id.*

2. Semi-Autonomous Production

When algorithms are designed to respond to data from external sources, they qualify as Semi-Autonomous.[23] These programs behave with a greater degree of freedom to produce statements that are not immediately intended by the programmer.[24] This can result in text that appears more natural and "human," which can be a desirable trait when interacting with external information.[25] This level of sophistication could also require less internal input and oversight, saving an operator's time and resources.[26]

One (in)famous example of Semi-Autonomous Production is Microsoft's AI chatbot, "Tay," for which Microsoft created an account on Twitter in 2016.[27] The program was designed to learn from external sources by interacting with other users on the platform, allowing it to appear more human.[28] Unfortunately, within a day of its debut, Tay's Twitter account began posting inflammatory and inappropriate statements based upon its interactions with other Twitter users.[29] The chatbot was quickly taken down by an embarrassed Microsoft, but the episode provides a significant warning about the dangers of allowing AI to generate and publish statements without oversight.[30]

A more familiar, everyday example of Semi-Autonomous Production is the autocomplete function available in search engines and word processors.[31] These functions are designed to respond to external user input and predict the next several words a user would like to type.[32] Like Tay, these programs take user input and extrapolate new statements to varying results: sometimes the statements produced by autocomplete are acceptable, and other times they can be problematic.[33]

There are few examples of Semi-Autonomous news stories which have made it to print. Two articles—one published in *The Guardian* in 2020 and one in *The New York Times* in 2021—were written using artificial intelligence

23. *See* Sears, *supra* note 15, at 1333-34.
24. *Id.*
25. *See id.*
26. Graefe, *supra* note 1.
27. John West, *Microsoft's Disastrous Tay Experiment Shows the Hidden Dangers of AI*, QUARTZ (July 21, 2022), https://qz.com/653084/microsofts-disastrous-tay-experiment-shows-the-hidden-dangers-of-ai/ [https://perma.cc/T4M5-NFQ7].
28. *Id.*
29. *Id.*
30. *See, e.g.*, *id.*
31. *See* Sears, *supra* note 15, at 1333-34.
32. Danny Sullivan, *How Google Autocomplete Works in Search*, GOOGLE SEARCH: THE KEYWORD (Apr. 20, 2018), https://blog.google/products/search/how-google-autocomplete-works-search/ [https://perma.cc/Y6AQ-AWJV].
33. Alex Hern, *(Auto)complete Fail: How Search Suggestions Keep Catching Google Out*, GUARDIAN (May 22, 2018, 8:09 AM), https://www.theguardian.com/technology/shortcuts/2018/may/22/autocomplete-fail-how-search-suggestions-keep-catching-google-out [https://perma.cc/DNF3-Q6N7].

to talk about artificial intelligence.[34] However, both of these articles required a good deal of editorial control over the algorithm in order to generate text that was suitable to print.[35] One editor noted that generating the article required producing eight different iterations and splicing them together,[36] while another pointed out that the algorithm took several tries because it kept getting stuck in an iterative loop.[37] If the goal of autonomous journalism is to require less user input while still generating seemingly natural statements, Semi-Autonomous Production may still have a long way to go.

So far, the question of liability for Semi-Autonomous Production has been averted through the application of Section 230 of the Communications Decency Act.[38] This Section provides in part that "No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider."[39] In other words, websites and content platforms cannot be held liable for information shared by third-party users. This is important because Semi-Autonomous Production is used most frequently by search engines and social media platforms.[40] These parties can argue that algorithmic statements occur because of third-party posts or links, meaning they cannot be held liable.[41]

As a result, cases involving liability for Semi-Autonomous Production have generally originated outside the United States:[42] a surgeon from Australia who sued Google for implying he was bankrupt through its autocorrect,[43] a former First Lady in Germany who sued because it implied she was a former escort,[44] and a Japanese man who sued Google for appending various crimes to his name when it was typed into the search bar.[45] As AI is

34. GPT-3, *supra* note 3; Kevin Roose, *A Robot Wrote This Book Review,* N.Y. TIMES (Nov. 21, 2021), https://www.nytimes.com/2021/11/21/books/review/the-age-of-ai-henry-kissinger-eric-schmidt-daniel-huttenlocher.html [https://perma.cc/ZSM2-YUZ7].

35. GPT-3, *supra* note 3 (editor's note describing how GPT-3 generated the article's text); Roose, *supra* note 32 (author's note describing how GPT-3 generated the text featured in the book review).

36. GPT-3, *supra* note 3 (editor's note describing how GPT-3 generated the article's text).

37. Roose, *supra* note 34 (author's note describing how GPT-3 generated the text featured in the book review).

38. Communications Decency Act of 1996, 47 U.S.C. § 230 (2016).

39. *Id.*

40. *See* Sears, *supra* note 15, at 1332-33.

41. Seema Ghatnekar*, Injury by Algorithm: A Look into Google's Liability for Defamatory Autocompleted Search Suggestions*, 33 LOY. L.A. ENT. L. REV. 171, 172 (2013).

42. *Id.* at 173-74.

43. Jeffrey P. Hermes, *Filing Lawsuits in the United States over Google Autocomplete Is . . .* , DIGIT. MEDIA L. PROJECT (Jan. 23, 2013, 5:03 PM), http://www.dmlp.org/blog/2013/filing-lawsuits-united-states-over-google-autocomplete [https://perma.cc/5YBR-A2BM].

44. *Google Auto-Correct Libellous, German Court Finds,* SYDNEY MORNING HERALD (May 15, 2013, 9:18 AM), https://www.smh.com.au/technology/google-autocorrect-libellous-german-court-finds-20130516-2jnfl.html [https://perma.cc/2VYN-EHQD].

45. Damien Gayle, *Google in Court After Man Complains Search Engine Automatically Adds Crimes After His Name,* DAILY MAIL (June 19, 2012, 2:25 PM), https://www.dailymail.co.uk/sciencetech/article-2161580/Google-court-man-complains-search-engine-automatically-adds-crimes-name.html [https://perma.cc/X7HE-796D].

more widely used by publishers rather than platforms, liability for defamation by algorithms may be extended to more defendants in the United States as well.

3. Fully Autonomous Production

The culmination of text-creating AI will be the fully autonomous production of speech.[46] This level of AI can create speech with little to no user input or intervention.[47] However, without a human overseer, a Fully Autonomous program could produce problematic statements that ultimately make it to publication. This category of AI is the least understood because it has not yet been fully realized.[48]

One specific risk associated with algorithmic speech is that false or defamatory statements produced by AI ultimately make it to print, leading to a publisher being sued for libel.[49] The concern has arisen in litigation but has not been directly addressed by American courts.[50] Another concern is the capacity for AI to be used in the spread of misinformation either intentionally or unintentionally.

4. Artificial Intelligence and Misinformation

Algorithmic speech may be a powerful tool for news organizations attempting to share legitimate news stories, but it may also become a weapon used in the propagation of disinformation.[51] Researchers have identified how advances in AI might be used to increase the effectiveness of disinformation campaigns by malicious actors.[52] Individuals who encounter false statements from these or other sources often have difficulty discerning that they are

46. *See* Sears, *supra* note 15, at 1333.
47. *Id.*
48. *See id.*
49. Lewis et al., *supra* note 6, at 65.
50. Ben Grubb, *Australian Doctor Withdraws Lawsuit Against Google*, EXAMINER (June 17, 2013), https://www.examiner.com.au/story/1579970/australian-doctor-withdraws-lawsuit-against-google/ [https://perma.cc/LRA3-MRGX]. After bringing a federal lawsuit in California, plaintiff Guy Hingston argued that Google should be held responsible for defamation by alleging he was bankrupt through its autocomplete feature. *Id.* When he typed his name into the search bar, the earliest options included "Guy Hingston bankrupt." *Id.* The case was never decided because the plaintiff withdrew the lawsuit. *Id.*
51. Cade Metz & Scott Blumenthal, *How A.I. Could be Weaponized to Spread Disinformation*, N.Y. TIMES (June 7, 2019), https://www.nytimes.com/interactive/2019/06/07/technology/ai-text-disinformation.html [https://perma.cc/EN5C-DNVE]; Giancarlo Mori, *The Rise of AI-Enabled Disinformation*, MEDIUM (May 7, 2021), https://gcmori.medium.com/the-rise-of-ai-enabled-disinformation-577e38fe724a [https://perma.cc/D2AN-SZSZ].
52. KATERINA SEDOVA ET AL., CTR. FOR SEC. & EMERGING TECH., AI AND THE FUTURE OF DISINFORMATION CAMPAIGNS 6 (2021), https://cset.georgetown.edu/wp-content/uploads/CSET-AI-and-the-Future-of-Disinformation-Campaigns.pdf [https://perma.cc/3UB4-Y93R].

untrue and may actively spread it further.[53] If an AI news aggregator or algorithmic speech program receives false or misleading statements without editorial safeguards, it may incorporate false statements into its news production.[54] This can create a significant problem for the truth-seeking public and for individuals who may be harmed by defamation.

The public's increased access to channels of communication through the Internet has compounded the problem of the potential spread of false information. Disinformation campaigns can use AI on social media particularly effectively because social media posts are usually short enough that it is difficult to distinguish between a human speaker and an algorithmic speaker.[55] Furthermore, after false or misleading statements are initially published, dissemination follows naturally as users post, repost, or share information with one another through various channels.[56]

As algorithmic speech increases in its use and sophistication, the threat of false or misleading statements also increases.[57] Inevitably, this misinformation will start to affect real individuals, causing reputational and other damage.[58] Legal and policy measures must be taken to ensure that the threat of reputational damage is kept to a minimum and that the public has access to trustworthy and reliable news, even from AI. One measure should be requiring publishers who use algorithmic speech for news production to exercise a reasonable duty of care in its journalistic process.

B. Libel and Defamation

Libel is a type of defamation, specifically "the publication of defamatory matter by written or printed words, by its embodiment in physical form or by any other form of communication that has the potentially harmful qualities characteristic of written or printed words."[59] Defamatory communication is that which "harm[s] the reputation of another as to lower him in the estimation of the community or to deter third persons from associating or dealing with him."[60] State courts generally follow the

53. Tom Buchanan, *Why Do People Spread False Information Online? The Effects of Message and Viewer Characteristics on Self-Reported Likelihood of Sharing Social Media Disinformation*, PLOS ONE 1 (Oct. 7, 2020), https://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0239666&type=printable [https://perma.cc/8YRB-K3T5].

54. John Villasenor, *How to Deal With AI-Enabled Disinformation*, BROOKINGS INST.: CTR. FOR TECH. INNOVATION (Nov. 23, 2020), https://www.brookings.edu/research/how-to-deal-with-ai-enabled-disinformation/ [https://perma.cc/UV7K-NZG8].

55. SEDOVA, *supra* note 52, at 5-6.

56. Buchanan, *supra* note 53, at 2-3; Villasenor, *supra* note 51.

57. *See generally* Saahil Desai, *Misinformation Is About to Get So Much Worse*, ATLANTIC (Sept. 27, 2021), https://www.theatlantic.com/technology/archive/2021/09/eric-schmidt-artificial-intelligence-misinformation/620218/ [https://perma.cc/D58A-SGR5].

58. Villasenor, *supra* note 54.

59. RESTATEMENT (SECOND) OF TORTS § 568 (AM. L. INST. 1977).

60. RESTATEMENT (SECOND) OF TORTS § 559 (AM. L. INST. 1977).

Restatement (Second) of Torts § 558[61] and recognize that a claim of libel requires:

> (a) false and defamatory statement concerning another;
> (b) an unprivileged publication to a third party;
> (c) fault amounting at least to negligence on the part of the publisher; and
> (d) either actionability of the statement irrespective of special harm or the existence of special harm caused by the publication.[62]

For purposes of analyzing claims with respect to autonomous journalism, this Note is most concerned with the third element of fault.

1. The Negligence Standard

The requirement of "fault amounting at least to negligence" describes the negligence standard as the minimum degree of fault required in a defamation claim.[63] For plaintiffs who are neither public officials nor public figures, the courts have left it to the individual states to determine the required degree of fault for these private figures to succeed on a claim of defamation, so long as they do not impose liability without fault.[64] The vast majority of states have declined to impose additional requirements on plaintiffs beyond negligence, so the negligence standard is generally applied to private individuals.[65] This standard requires that a plaintiff prove that, in addition to a publication being false, the defendant knew it to be false *or* lacked reasonable evidence to believe it was true *or* acted negligently in failing to ascertain its truth.[66] It imposes on defendants a duty of reasonable care in verifying the truth or falsity of information published about private individuals. Publishers must be justified in the belief that their publications are true.[67]

In practice, defendants seeking to prove they fulfilled the duty of reasonable care can rely on various types of evidence. Juries may be instructed to consider "the reliability, the nature of the sources of the defendant's information, its acceptance or rejection of the sources, and its care in checking upon assertions."[68] The amount of urgency in reporting a particular story, the need to investigate a matter thoroughly, and whether

---

61. *See, e.g.*, McAdoo v. Diaz, 884 P.2d 1385 (Alaska 1994); Boswell v. Phoenix Newspapers, Inc., 730 P.2d 178 (Ariz. Ct. App. 1985).

62. RESTATEMENT (SECOND) OF TORTS § 558 (AM. L. INST. 1977).

63. *Id.*

64. *See* Gertz v. Robert Welch, Inc., 418 U.S. 323, 347-48 (1974).

65. *See* 99 AM. JURIS. PROOF OF FACTS, 3D *Proof of Facts Establishing Affirmative Defenses Against a Claim for Defamation* § 17 (2008).

66. *See* Gazette, Inc. v. Harris, 229 Va. 1, 15 (1985); *see also* RESTATEMENT (SECOND) OF TORTS § 580B (AM. L. INST. 1977).

67. *See Harris*, 229 Va. at 5; *see also* Carney v. Santa Cruz Women Against Rape, 221 Cal. App. 3d 1009, 1016 (Ct. App. 1990); *see also* Duchesnaye v. Munro Enters., Inc., 125 N.H. 244, 251 (1984).

68. Curtis Pub. Co. v. Butts, 388 U.S. 130, 156 (1967).

independent efforts were taken to corroborate information may also play a role.[69] Affidavits from expert journalists attesting that publishers acted in accordance with professional journalistic standards may also provide a strong defense because these standards are well known and because courts frequently recognize them.[70] These procedures include being thorough and fair, carefully attributing sources and quotes, not phrasing statements in a way to create unsupported implications, relying on multiple sources, and giving news subjects an opportunity to respond or comment.[71] When publishers depart from standard procedure in fact-checking information, they risk breaching the reasonable duty of care required of journalists under the negligence standard.

In *Gertz v. Robert Welch, Inc.*, the Supreme Court established the negligence standard as the minimum degree of fault private individuals must prove to succeed in a libel claim, abolishing the rule of strict liability for defamation.[72] This negligence standard is less demanding to publishers than the strict liability standard, which held defendants liable for *any* false information they published, making one's only defense the truthfulness of the statement.[73] Under the negligence standard, publishers can successfully defend themselves by showing they exercised reasonable care and published the false information without knowing it was false.[74] However, the negligence standard is not as demanding as the actual malice standard, which requires more fault on the part of publishers when discussing public officials and public figures.[75]

2. The Actual Malice Standard

In 1967, the U.S. Supreme Court ruled in *New York Times v. Sullivan* that for a printed statement about a public official to be considered libelous, the public official must show that a statement not only fulfills the original four elements of defamation (including fault amounting at least to negligence), but that the statement was also made with "actual malice."[76] The Court in that case defined actual malice as "knowledge that [a statement] was false or [made] with reckless disregard of whether it was false or not."[77] Shortly thereafter, the Court extended this standard to public figures or those about whom the public has a justified interest.[78]

---

69. *Id.* at 157-59.
70. Greenberg v. CBS Inc., 69 A.D.2d 693, 709-10 (N.Y. App. Div. 1979).
71. *Practical Tips for Avoiding Liability Associated with Harms to Reputation*, DIGIT. MEDIA L. PROJECT, https://www.dmlp.org/legal-guide/practical-tips-avoiding-liability-associated-harms-reputation [https://perma.cc/PL72-ZA3T] (last visited Sept. 12, 2022).
72. *Gertz*, 418 U.S. at 345-48.
73. *Id.* at 340-41.
74. *Id.* at 334.
75. *Id.* at 342.
76. *See* New York Times Co. v. Sullivan, 376 U.S. 254, 279-80 (1964).
77. *Id.* at 280.
78. *See* Curtis Pub. Co. v. Butts, 388 U.S. 130, 158 (1967). The Court also identified "public figures" as those who are "involved in issues in which the public has a justified and important interest." *Id.* at 134. In this case, a prominent university football coach. *Id.*

The purpose of this heightened standard is to protect the defendant's right to freedom of expression when speaking about public individuals and matters of "the highest public interest and concern."[79] As noted by the Court, citizens will likely have an interest in speaking, often critically, about public figures and officials.[80] Applying the less stringent negligence standard to libel claims may discourage free discussion and make citizens unwilling to speak out on public matters, which would be antithetical to the purpose of the First Amendment's freedom of speech.[81]

Another reason courts have cited for allowing the actual malice standard is that public figures and public officials have greater access to news media and resources for making public statements.[82] Given these resources, public officials and public figures have a greater opportunity to set the record straight if a defamatory statement is widely publicized.[83]

Proving the existence of actual malice presents an obstacle to plaintiffs, even in cases involving human authors. The Supreme Court has provided for the use of direct or circumstantial evidence, including threats, prior or subsequent statements of the defendant, evidence indicating a rivalry or hostility, and other facts showing a reckless disregard of a plaintiff's rights.[84] Malice in this case speaks to a publication's intent or motive, specifically the publisher's "ill will, spite, hatred and an intent to injure."[85] It can be difficult to prove the internal motivations of a particular party in the best of circumstances, which is why circumstantial evidence is permitted in such cases.[86] However, this creates an even bigger problem in cases involving algorithmic speech. Because speech is produced mechanically, it could be impossible for plaintiffs to demonstrate that a statement's "author" either had serious doubts about what it was saying or harbored ill will towards the subject.[87] In order to protect individuals from algorithmic defamation, and the public from misinformation, courts should modify the requirements for liability in cases involving AI.

Included in the definition of actual malice is a "wanton or reckless indifference or culpable negligence."[88] This addition to the standard can be confusing when distinguishing the actual malice standard from the negligence standard. The Court clarified its position in *St. Amant v. Thompson*, admitting that reckless disregard "cannot be fully encompassed in one infallible definition" but that requirements include "sufficient evidence to permit the conclusion that the defendant in fact entertained serious doubts as to the truth of the publication."[89] Including "reckless disregard" for the truth within the

79. *Sullivan*, 376 U.S. at 266.
80. *See id.* at 270-71.
81. *See id.* at 271-72, 296-97 (Black, J., concurring).
82. *Id.* at 304-05 (Goldberg, J., concurring); *Butts*, 388 U.S. at 155.
83. *Butts*, 388 U.S. at 155.
84. Herbert v. Lando, 441 U.S. 153, 164 n.12 (1979) (quoting 50 AM. JURIS. 2D *Libel and Slander* § 455 (1970)).
85. *Id.* at 162 (quoting *Butts*, 388 U.S. at 138 n.3).
86. *Id.* at 172-73.
87. Lewis et al., *supra* note 6, at 69.
88. *Herbert*, 441 U.S. at 162.
89. St. Amant v. Thompson, 390 U.S. 727, 730-31 (1968).

definition of actual malice does not impose a reasonable duty to check the reliability of statements published about public officials and figures.[90]

3. Who Can Be Liable?

Traditionally, the original author or publisher of defamatory information is the individual or company liable for any injury it causes.[91] Under the Doctrine of Republication, the original author is not liable for its republication by a third party if they did not authorize, or could not have reasonably foreseen, its republication.[92] In that case, the party that repeats or republishes untrue statements can be held liable.[93] Under the negligence standard, this essentially imposes a duty on "republishers" to fact-check the original information or risk incurring liability.

An important exception to the Doctrine of Republication is the "wire service defense," which allows the media to republish defamatory statements without liability in some circumstances.[94] The rule was first used in *Layne v. Tribune Co.*, where a defendant newspaper company republished a libelous story about the plaintiff that it received by wire.[95] The court ruled that because the original source was a "generally recognized reliable source of daily news" there was no defamation unless the newspaper acted recklessly or carelessly in reproducing the story.[96] The defense developed a standard of reasonable duty of care, allowing smaller news outlets to share news from around the country without fear of liability, so long as they read the original article and did not detect any reason to doubt its truthfulness.[97] The existence and nature of the wire service defense differs state by state, but states that recognize it generally require that: (1) a publisher received the news from a reputable news agency; (2) they did not know the information was false; (3) the news item does not indicate any reason to doubt its veracity; and (4) the publisher does not substantially alter the news items when republishing it.[98]

Ultimately, the tort of libel allows for people damaged by the untrue words of another to recover for damage to their reputations. It also provides powerful incentives to those who publish to exercise care that they are sharing information which is correct and does not infringe on a person's right to privacy. The provisions in the Restatements provide for an injured plaintiff to recover from the actor most responsible for an injury done to them.[99] However, if an injured party cannot recover because of the requirements

90. *Id.* at 731.
91. 53 C.J.S. *Libel and Slander* § 91 (1948).
92. 53 C.J.S. *Libel and Slander* § 102 (1948).
93. RESTATEMENT (SECOND) OF TORTS § 559 (AM. L. INST. 1977).
94. Jennifer L. Del Medico, Comment, *Are Talebearers Really as Bad as Talemakers?: Rethinking Republisher Liability in an Information Age*, 31 FORDHAM URB. L.J. 1409, 1410 (2004).
95. Layne v. Tribune Co., 146 So. 234, 235-36 (Fla. 1933).
96. *Id.* at 238.
97. Del Medico, *supra* note 94 at 1412-13.
98. *Wire Service Defense*, DIGIT. MEDIA L. PROJECT, https://www.dmlp.org/legal-guide/wire-service-defense [https://perma.cc/3HSH-LCMG] (last visited Sept. 12, 2022).
99. RESTATEMENT (SECOND) OF TORTS § 578 (AM. L. INST. 1977).

imposed by law, courts should reconsider those standards' purpose and effectiveness. The following analysis will consider how AI interacts with the actual malice standard and argue that the standard is insufficient to ensure that the purposes of libel are met.

III. ANALYSIS

Scholars in journalism and communications law have identified news organizations' growing concern for inadvertently spreading libel through artificial intelligence. In 2018, scholars writing for Journalism and Mass Communication Quarterly saw "Libel by Algorithm" as a potential legal hazard that journalists should be wary of in the near future.[100] They outlined several situations in which algorithms have played a part in spreading disinformation, summarized the scholarship surrounding whether First Amendment protection should be given to algorithms, and pointed out that public individuals who are plaintiffs would be unlikely to recover unless the court is willing to create a new standard of liability.[101] One point of concern is the difficulty of showing actual malice on the part of AI users, who may not understand the algorithmic speech creation process.[102] Other scholars have corroborated these concerns, pointing out prior cases that suggest libel via algorithmic speech is possible, and they assert the difficulty of successfully bringing a claim for defamation against AI under the actual malice standard.[103]

Developments in communications technology have given people the ability to publish and share information on a larger scale than ever. This has led to a bounty of information being freely available to individuals across the world, but this also makes identifying the source of information, and its truthfulness, far more difficult.[104] Algorithms in particular have little ability to verify the truthfulness of statements and can easily republish or redistribute the libelous words of others by mistake.[105] This greater risk merits imposing a greater responsibility on AI users to verify the speech it produces.

The existing standard requiring actual malice for public figures to bring a claim of defamation does not sufficiently impose this responsibility. Artificial intelligence itself does not engage in the subjective decision-making process evaluated under the actual malice standard.[106] Actual malice requires a plaintiff to demonstrate that the author or publisher possessed ill will

100. Lewis et al., *supra* note 6, at 60-81.

101. *Id.* at 63-69.

102. *Id.* at 69.

103. Ghatnekar, *supra* note 41, at 173-74; *see also* Peter Georgiev, *A Robot Commits Libel. Who Is Responsible?*, REYNOLDS JOURNALISM INST. (Feb. 20 2019), https://rjionline.org/news/a-robot-commits-libel-who-is-responsible/ [https://perma.cc/3VKL-4PZT].

104. Filippo Menczer & Thomas Hills, *Information Overload Helps Fake News Spread, and Social Media Knows It*, SCI. AM. (Dec. 1, 2020), https://www.scientificamerican.com/article/information-overload-helps-fake-news-spread-and-social-media-knows-it/ [https://perma.cc/R42W-6VFG].

105. *Id.*

106. Graefe, *supra* note 1.

towards the plaintiff or had serious doubts about the veracity of a defamatory statement.[107] However, algorithms are designed to produce information mechanically, and it would be impossible to prove they possessed ill will or doubts in the traditional sense.[108] Thus, negligence is a better standard for considering claims of defamation.

A. Applying the Negligence Standard

The best way to analyze the advantages of applying the negligence standard over the actual malice standard for cases involving algorithmic speech is through illustration. Consider the following scenario:

A tech company releases a chatbot named ALICE that is designed to interact with users on social media. ALICE can create short articles about the user's local weather, local news, and current events. ALICE is programmed to learn from the language of human users on the platform and produce statements that are calculated to foster the greatest amount of engagement with the online community. ALICE is also programmed to avoid making controversial statements or commenting on heated issues, as identified by its developers. However, despite this safeguard, ALICE engages in speech with several users about a small-town politician. Based on the information in her interactions, ALICE goes on to make statements to a large number of other users that strongly imply the politician is involved with organized crime. The users, supposing that these are news announcements, take her statements at face value. There is no evidence that the politician has connections with organized crime, but she suffers reputational damage regardless. She sues the tech company for defamation.

If ALICE were a human author ("Alice") producing news and statements for social media, her statements would be reviewed under the actual malice standard because the plaintiff is a public figure.[109] Under these circumstances, inquiries into Alice's journalistic process, state of mind, and her own personal knowledge would center around whether she knew that her statements were false, if she behaved with reckless disregard as to whether they were false, or if she bore ill will or an intent to injure the politician.[110] Alice could be questioned and cross-examined, and the company's policies regarding its journalists could be used as evidence in convincing a jury that there was or was not actual malice.[111]

However, in this scenario, ALICE is an algorithm. Therefore, applying the actual malice standard would yield an incoherent analysis. Inquiries into ALICE's "journalistic process" would yield little insight into whether ALICE "believed" her statements to be true.[112] Similarly, it would be difficult, if not impossible, to show that ALICE bore ill will or resentment towards any

107. *See* New York Times Co. v. Sullivan, 376 U.S. 254, 283-84 (1964).

108. Lewis et al., *supra* note 6, at 69.

109. *See Sullivan*, 376 U.S. at 279-80.

110. *See id.* at 279-80, 283-84.

111. Herbert v. Lando, 441 U.S. 153, 164 n.12 (1979) (quoting 50 AM. JURIS. 2D *Libel and Slander* § 455 (1970)).

112. Lewis et al., *supra* note 6, at 68.

individual.[113] Inquiries into ALICE's programmers—those who created the program—and operators—those who used the program to produce statements—would also be stymied; programmers may not even be aware of the politician's identity, and operators presumed that the program was operating properly.[114] At best, the plaintiff could try to make a case for wanton and reckless indifference on the operators' part concerning whether published statements were true or not.[115] In all likelihood, the damage done to the public figure and her reputation would remain unresolved, and there would still be a risk of spreading misinformation.

However, if a court were to apply the negligence standard to the Scenario, the analysis is much more coherent. The question before the court would be whether ALICE and her handlers fulfilled a reasonable duty of care to determine the truth of her statements.[116] A jury could be directed to consider the reliability of ALICE's sources of information, her acceptance or rejection of particular sources, and the algorithm's methods of checking upon assertions.[117] Programmers and operators could testify about the nature of ALICE's fact-checking method, and whether she derives information from any common profile or if she corroborates stories with reliable sources. Experts could testify regarding whether the algorithm's methods meet standards of journalistic procedure. Under this standard, a defendant would still prevail if they were to show that ALICE's safeguards and methods are reasonable enough to fulfill the duty of care.[118] However, the plaintiff in this situation also has an opportunity to succeed if she demonstrates ALICE's programming and publisher's procedures lead to negligent, untrustworthy statements.[119]

B. Libel Defendants in cases involving Artificial Intelligence

The likely defendants to a claim of defamation involving artificial intelligence are the algorithm's operators (news organizations and individuals who use AI to produce statements and publish them) and creators (programmers and software developers that create the speech-producing AI).[120] In lieu of demonstrating malice on the part of the algorithm, showing malice on the part of human operators or creators would satisfy the standard.[121] However, in many cases, it may be too difficult to show that these human actors demonstrated malice.[122] Either might argue they justifiably relied on AI tools, or they might merely assert that no duty to verify

113. *Id.*
114. *Id.*
115. *Herbert*, 441 U.S. at 162.
116. RESTATEMENT (SECOND) OF TORTS § 580B (AM. L. INST. 1977).
117. *See, e.g.*, 16B AM. JURIS. PLEADING & PRAC. FORMS *Libel and Slander* § 286 (last updated July 2022).
118. *See* RESTATEMENT (SECOND) OF TORTS § 580B cmt. d. (AM. L. INST. 1977).
119. *See id.*
120. *See* Lewis, supra note 6, at 71-72.
121. *See id.* at 68.
122. *See generally id.* at 69.

information exists under the standard outlined in *St. Amant v. Thompson*.[123] Because of this impossible standard, adopting a negligence standard would be more appropriate for cases involving autonomous journalism.

Publishers are generally the less sophisticated of the two groups when it comes to understanding the risks of utilizing an algorithmic speech program.[124] Fortunately, news organizations often have procedures in place to facilitate accurate reporting.[125] However, some publishers may allow statements produced by AI to be printed without editorial review.[126] Unless courts place an affirmative duty to review statements produced by AI, operators may avoid responsibility in cases involving public officials by citing their lack of serious doubts about the information algorithms produce. Ultimately, actual malice could become an impossible standard to prove, and AI in the wrong hands would become a tool of blatant disinformation.

Under the negligence standard, operators would have a reasonable duty of care to seek accurate reporting.[127] This is the standard to which individuals and news organizations are already held when publishing statements about private individuals, so it would merely require extending the same care to the public when producing articles autonomously.[128]

Liability for software developers is a rapidly developing field, and it has only barely touched the legal topic of defamation.[129] However, software developers can be held liable for negligent design where a defect in software causes physical injury, fails to protect private information, or possesses another design defect.[130] Under a theory of negligence in software development, the plaintiff must show that (1) the developer had a duty to provide functioning software; (2) the developer breached this duty; (3) the user suffered harm; and (4) the harm was caused by the software.[131] For defamation, this is the most likely route by which developers may be held liable for the actions of their programs.

Extending the negligence standard for libel to software developers would utilize the existing framework for negligence. Developers are liable for software that does not operate correctly if they were negligent in its production. Like other news producers and original authors, developers should have a duty to ensure algorithmic speech software was reasonably programmed to produce statements which are true. This could include implementing fact-checking software, requiring statements to include corroborating sources, or flagging potentially sensitive statements for human review. If an algorithm does not fulfill this duty and routinely produces statements that are false and injure another person, the software breaches this

123. *See* St. Amant v. Thompson, 390 U.S. 727, 730-31 (1968).
124. Graefe, *supra* note 1.
125. *See* Lewis et al., *supra* note 6, at 69-70.
126. *Id.*
127. *See Thompson*, 390 U.S. at 734 (Fortas, J., dissenting).
128. *See* Gertz v. Robert Welch, Inc., 418 U.S. 323, 347-48 (1974).
129. *See* 46 AM. JURIS. TRIALS 687 *Failure of Performance in Computer Sales and Leases* §§ 1, 17 (1993).
130. *See* 46 AM. JURIS. TRIALS 687 *Failure of Performance in Computer Sales and Leases* § 24 (1993).
131. *See id.*

duty, and the developer could be liable.[132] This would encourage developers to design algorithmic speech in a way that does not produce libelous statements.

However, to more effectively protect themselves from this kind of liability, software companies have increasingly relied on indemnification clauses in license agreements.[133] Under one of these agreements, a vendor agrees to license their software to another entity in exchange for payment, but the vendor often includes language that seeks to limit the software developer's liability for injuries caused by the software.[134] In the case of AI, it is easy to imagine a situation where creators license their algorithmic speech programs to publishers under a license which limits their liability for defamation. In such a situation, they would more effectively shield themselves from responsibility for defamation claims but shift the burden of liability solely onto the operator for use of their software.

Under the actual malice standard, such license agreements would block any claim for defamation by public figures or public officials. Users could defend themselves on the grounds that they possessed no actual malice and trusted in the software to produce correct statements, while software developers would use indemnification clauses to avoid responsibility. Under the negligence standard, however, developers and users would be encouraged to make clear in the terms of their agreements who is responsible for fact-checking and which parties are responsible for potentially defamatory statements produced by algorithmic speech programs.

Adopting the negligence standard in all cases of algorithmic speech would certainly make it more difficult for those implementing it to dodge responsibility for producing libelous statements. However, there are rational reasons why the actual malice standard is used for plaintiffs who are public individuals in the first place, and there are reasonable concerns with abandoning that standard with respect to AI.

C. Concerns with the Negligence Standard

Opponents to adopting the negligence standard for algorithmic speech regarding public figures may cite several concerns. The greatest of these is that it may restrict freedom of speech. Courts have been unwilling to restrict First Amendment rights even for nontraditional speakers,[135] but the non-personhood of algorithms and the reduced human control over algorithmic speech may warrant reconsideration. Proponents of the actual malice standard may also argue that it is justified given the privileged access that public officials and figures have to certain channels of communication. However, these individuals' ability to counter disinformation has been diminished by AI, while private individuals' access to mass communication channels has

132. *See id.*

133. *See, e.g.*, Peter M. Moldave, *Software Agreements*, *in* DRAFTING AND NEGOTIATING MASSACHUSETTS CONTRACTS § 13.8 (John F. Cohan ed., 2022).

134. *See id.*

135. *See* Citizens United v. FEC, 558 U.S. 310, 372-73 (2010).

increased, closing the gap between the two groups in their ability to counter false information about themselves.[136]

1. Freedom of Speech

The First Amendment does not sanction a statement of libel or defamation, nor does it remove civil liability from those who participate in it.[137] However, courts have recognized concerns that the threat of a defamation claim may stifle the freedom of expression which the First Amendment is meant to protect.[138] The original purpose of the actual malice standard was to protect individuals' freedom to speak critically about public officials and figures.[139] Applying the less permissive negligence standard may impose a burden on publishers whenever they want to criticize those in power. This could ultimately discourage editorial journalism or the free flow of information and opinions among the public.[140]

However, the existence of automated journalism raises the question of whether algorithmic authors deserve the same free speech rights as living individuals.[141] Human persons' freedom of speech under the actual malice standard would not be curtailed by applying the negligence standard to AI because living individuals can testify to their knowledge or ignorance of the truthfulness of their own statements.

Despite obvious differences between living individuals and non-human speakers, in the United States, many courts have been unwilling to restrict freedom of speech even for non-traditional speakers.[142] In *Citizens United v. FEC*, the Supreme Court upheld a corporation's right to political speech on First Amendment grounds.[143] In his concurrence, Justice Scalia emphasized that the first Amendment "is written in terms of 'speech,' not speakers." [144]

However, algorithmic speech can be distinguished from corporate speech on several grounds. Corporations and business entities represent groups of individual humans. One could argue that they only qualify for freedom of speech under the First Amendment because corporate "persons" are merely legal stand-ins for groups of people.[145] While this ground for granting personhood has been attacked by critics pointing out the nature of control of corporations,[146] even this defense does not apply as strongly to algorithmic speech. Groups of individual humans are involved in the

136. *See* Menczer & Hills, *supra* note 104.
137. *See* New York Times Co. v. Sullivan, 376 U.S. 254, 264-65 (1964).
138. *See id.* at 266.
139. *See id.* at 270.
140. *See id.*
141. Toni M. Massaro et. al., *Siri-ously 2.0: What Artificial Intelligence Reveals About the First Amendment*, 101 MINN. L. REV. 2481, 2506 (2017).
142. *See* Massaro & Norton, *supra* note 10, at 1183-85.
143. Citizens United v. FEC, 558 U.S. 310, 372 (2010).
144. *Id.* at 392-93 (Scalia, J., concurring).
145. *See* Tamara R. Piety, *Why Personhood Matters*, 30 CONST. COMMENT. 361, 370-71 (2015).
146. *See* Nadia Imtanes, *Should Corporations Be Entitled to the Same First Amendment Protections as People?*, 39 W. ST. U. L. REV. 203, 214 (2012).

programming and calibration of AI, but the ultimate speech product is not directly produced by human input.

Human control over speech is another distinguishing factor between algorithmic speech and corporate speech. Political contributions and corporate statements are decided and controlled by individual humans.[147] These individuals may represent a small group of a corporate body, but all actions are ultimately decided by humans.[148] For algorithmic speech, human control of speech is sacrificed to one degree or another for the benefit of efficient production.[149] One need only consider the example of Tay AI to recall that generated speech can stray far from its intended purpose and quickly get out of control.[150] The closer algorithmic speech gets to fully autonomous production, the further it gets from the control of human persons.

Furthermore, many of the criticisms leveled at extending First Amendment protections and personhood to corporations also apply to extending them to AI. The foremost criticism is that our society has a "philosophical, political, and moral commitment to the equality of human beings under the law" that we do not extend to fictional persons.[151] For many people, their basic instinct is to distinguish between the rights extended to living humans and fictional persons. Another argument against extending freedom of speech to fictional persons is that it seriously limits the power of the government to regulate in the public interest.[152]

Both of these arguments may apply to algorithmic speech just as powerfully as they do to corporate speech. According to one survey, an overwhelming majority of Americans believe that AI should be carefully managed.[153] Among the highest concerns of those surveyed was the need to prevent AI from violating privacy and civil liberties, and to prevent the spread of fake and harmful content online.[154] These results suggest a fundamental understanding of the need for government regulation and a distinction between the rights and privileges of human beings as opposed to artificial entities.

There are fundamental differences between human beings with freedom of speech—the foundation for maintaining the actual malice standard—and AI. Due to these differences, algorithmic speech produced by AI does not require the same protection of the actual malice standard, and the negligence standard should be applied to statements produced by machines instead.

147. *See* Piety, *supra* note 145, at 372-73.
148. *Id.*
149. *See* Sears, *supra* note 15, at 1333-34.
150. West, *supra* note 27.
151. Piety, *supra* note 145, at 385.
152. *Id.* at 387.
153. Baobao Zhang & Allan Dafoe, *Artificial Intelligence: American Attitudes and Trends*, CTR. FOR GOVERNANCE OF AI 10 (2019), https://governanceai.github.io/US-Public-Opinion-Report-Jan-2019/us_public_opinion_report_jan_2019.pdf [https://perma.cc/9FJL-F8N2].
154. *Id.* at 3-4.

2. Channels of Effective Communication

One reason for allowing more permissive speech on the part of reporters with regards to public figures and public officials is that they "enjoy significantly greater access to the channels of effective communication and hence have a more realistic opportunity to counteract false statements than private individuals normally enjoy."[155] However, with the rise of digital communications technology and the development of AI, public officials and figures' ability to counteract disinformation has been impaired, while their relative advantage over private individuals has diminished. Given this effect of expanding AI, the Court should reevaluate distinguishing between private and public subjects of defamation where AI is concerned.

The expansion of social media, algorithms that drive engagement, and the glut of information and disinformation available online contribute to how difficult it is to counter false statements.[156] Researchers have found that the overload of information available on the internet has contributed to individuals selecting sources which confirm their own biases.[157] Furthermore, a large percentage of adults in the United States rely on social media to get news.[158] This can lead people to rely on information from bots, automated social media accounts that impersonate humans, which are often designed to share disinformation.[159] This, combined with diminishing trust in traditional media,[160] has significantly impaired the effectiveness of public officials' resources in combating false statements.

Meanwhile, recent developments in communications technology and the social media landscape have granted the resources needed to disseminate information to more individuals. Around seven in ten Americans use social media to connect, read news, share information, and enjoy themselves.[161] It is such an effective way of communicating that nearly all members of Congress use social media to communicate with the public,[162] and lately presidents have

155. Gertz v. Robert Welch, Inc., 418 U.S. 323, 344 (1974).

156. Menczer & Hills, *supra* note 104.

157. *Id.*

158. Mason Walker & Katerina Eva Matsa, *News Consumption Across Social Media in 2021*, PEW RSCH. CTR. 3-4 (Sept. 20, 2021), https://www.pewresearch.org/journalism/wp-content/uploads/sites/8/2021/09/PJ_2021.09.20_News-and-Social-Media_FINAL.pdf [https://perma.cc/MRT6-533X].

159. *See generally How Is Fake News Spread? Bots, People Like You, Trolls and Microtargeting*, CTR. FOR INFO. TECH. & SOC. (2022), https://www.cits.ucsb.edu/fake-news/spread [https://perma.cc/74CW-S4XU].

160. Megan Brenan, *Americans' Trust in Media Dips to Second Lowest on Record*, GALLUP (Oct. 7, 2021), https://news.gallup.com/poll/355526/americans-trust-media-dips-second-lowest-record.aspx [https://perma.cc/ZE5L-LNV5].

161. *Social Media Fact Sheet*, PEW RSCH. CTR. (Apr. 7, 2021), https://www.pewresearch.org/internet/fact-sheet/social-media/ [https://perma.cc/EXQ2-RT4Q].

162. Kendra Kumor et al., *Improving Communication with Public Officials on Social Media: Proposals for Protecting Social Media Users' First Amendment Rights*, FORDHAM UNIV. SCH. L.: DEMOCRACY AND THE CONST. CLINIC 3-4 (2021), https://www.fordham.edu/download/downloads/id/15275/improving_communication_with_public_officials_on_social_media.pdf [https://perma.cc/Q997-J94B].

used social media as an effective means of communication.[163] These resources are widely available to Americans. This is not to say that there is no distinction between private and public individuals but demonstrates that the gap between public figures' and the majority of citizens' ability to reach large numbers of people is closing.

The expansion of channels of effective communication makes eliminating the actual malice standard appropriate specifically for AI because these new channels are how algorithmic speech can do the most harm. Public figures' ability to counter disinformation is diminished to the extent that AI is used to target willing recipients and amplify distrust in the individuals who benefit from the actual malice standard.[164] Even with their remaining advantages, AI will alter the landscape so dramatically that public officials may need legal protection which is currently unavailable.

IV. CONCLUSION

AI will create unique opportunities and advantages in the field of journalism as technology becomes more autonomous and sophisticated. It has already provided significant advantages by reducing the time and resources required to report stories that are largely "by-the-numbers," and it promises to become a useful tool in stories that are more nuanced and editorial in nature.[165] However, it has also led to some missteps which reveal the dangers of relinquishing editorial control to an algorithm and allowing programs to publish statements on their own.[166] Without proper editorial oversight, fully autonomous journalism risks propagating false and damaging statements about individuals.[167]

Algorithmic speech cannot be shown to be a product of actual malice the same way that human speech can.[168] Algorithms produce speech mechanically according to their programming, and it cannot be demonstrated that they doubt or believe information that they produce.[169] Finding actual malice on the AI's operators or creators' part is also difficult, as they will only be required to show a lack of serious doubts in the statements of the program, regardless of the harm caused.[170]

The negligence standard is better suited to addressing the concerns of defamation authored by artificial intelligence, even in the case of public officials and public figures. The negligence standard imposes a duty of reasonable care on publishers to check the truthfulness of its statements about

163. Daniel Victor, *When Joe Biden Took the White House, He Also Took @WhiteHouse*, N.Y. TIMES (Jan. 20, 2021), https://www.nytimes.com/2021/01/19/technology/biden-white-house-twitter-account.html [https://perma.cc/DLC6-VCDC].

164. SEDOVA, *supra* note 52, at 5.

165. Underwood, *supra* note 2.

166. West, *supra* note 27.

167. Lewis et al., *supra* note 6, at 65.

168. *Id.*

169. *Id.*; *see* Sears, *supra* note 15, at 1333-34.

170. Lewis et al., *supra* note 6, at 66.

individuals.[171] This duty is not only desirable with respect to checking statements generated by AI as a matter of policy, but it is essential in ensuring that the technology is used responsibly, in a way that does not contribute to disinformation and the erosion of the public's access to accurate information. The negligence standard can be applied effectively to both operators and developers of algorithmic speech technology.

Given the unique nature of algorithmic speech and its potential role in journalism, courts should not hesitate to adapt the standards for defamation as they apply to AI for all categories of individuals. Courts should apply the negligence standard when evaluating claims for libel or defamation of public individuals when false statements are generated by AI. By implementing this standard, human actors involved in publishing defamatory statements produced autonomously would appropriately bear the burden of ensuring those statements are accurate and that these powerful new technologies are implemented responsibly.

171. Gertz v. Robert Welch, Inc., 418 U.S. 323, 347-48 (1974).

Here, There, and Everywhere: Defining the Boundaries of the "Schoolhouse Gate" in the Era of Virtual Learning

Robin Briendel*

TABLE OF CONTENTS

* J.D., May 2023, The George Washington University Law School; B.A., Politics & Education Studies, Brandeis University. I would like to thank the entire FCLJ staff for their hard work in producing this publication. A special thanks to Sarah Morris, Journal Adjunct, and Andrew Seneviratne, Notes Editor, for their encouragement and guidance. Finally, I would like to thank my family for their unconditional love and support throughout the note writing process.

I. INTRODUCTION

It's 3:00 PM on a Thursday. The last school bell of the day rings at XYZ High School, signaling the end of the school day. A group of friends leave their algebra class and walk to Starbucks. On their walk, the friends discuss a classmate of theirs whom they dislike, another ninth grader—Student A. During the conversation, the girls refer to Student A as "fat," "ugly," and "stupid." One of the students in the group, Student B, creates a meme in which she superimposes Student A's Facebook profile picture on an image of Fiona, the ogre from the movie *Shrek*, with the caption "Weird Fat Fugly Ogre." Student B posts the meme on Twitter and shares it with her friends. Her friends retweet the meme and send it to additional students who are still at school waiting for soccer practice to begin. In only a few hours, the meme is circulated to much of the student body of XYZ High School. By midnight, it has been retweeted 350 times, has 1,500 likes, and has 200 comments.

Too afraid to face her peers, Student A refuses to go to school the following day. Enraged, her mother drives to the school with printed copies of the offending tweet and demands a meeting with the principal. Following the meeting, the principal identifies Student B as the meme's creator. He calls Student B to his office and suspends her from school for ten days for bullying Student A.

Weeks go by, and Student A remains distraught. Recognizing signs that her daughter, Student A, has started excessively exercising and restricting her calorie intake, Student A's mother enrolls her in an eating disorder program for teenagers affiliated with a local hospital. Around the same time, Student B, who realizes that her suspension will reflect poorly upon her as she applies to college, sues the school district and the principal, arguing that her suspension was an unconstitutional infringement of her First Amendment right to free speech.

While this might seem farfetched to some, this anecdote is based upon an amalgamation of lower court cases,[1] court documents,[2] and recent news stories.[3] Since the first social media website was introduced to the public in

1. *See* J.C. *ex rel.* R.C. v. Beverly Hills Unified Sch. Dist., 711 F. Supp. 2d 1094, 1098 (C.D. Cal. 2010); *see also* A.S. *ex rel.* Schaefer v. Lincoln Cnty. R-III Sch. Dist., 429 F. Supp. 3d 659, 664 (E.D. Mo. 2019).

2. *See* Reply Brief for Petitioner at *2-3, Mahanoy Area Sch. Dist. v. B.L. *ex rel.* Levy, 141 S. Ct. 2038 (2021) (No. 20-255), 2021 WL 1549729; *see also Mahanoy*, 141 S. Ct. at 2062-63 (Thomas, J., dissenting).

3. *See, e.g.*, Monica Anderson et al., *A Majority of Teens Have Experienced Some Form of Cyberbullying*, PEW RES. CTR. 2-3 (Sept. 27, 2018), https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2018/09/PI_2018.09.27_teens-and-cyberbullying_FINAL.pdf [https://perma.cc/Y5LS-QGGV]; *see also* Georgia Wells et al., *Facebook Knows Instagram Is Toxic for Teen Girls, Company Documents Show*, WALL ST. J. (Sept. 14, 2021, 7:59 AM), https://www.wsj.com/articles/facebook-knows-instagram-is-toxic-for-teen-girls-company-documents-show-11631620739 [https://perma.cc/TUD7-M3CA].

1997,[4] at least fifty federal court cases have been brought by students challenging the constitutionality of disciplinary measures taken against them by their schools for their off-campus speech.[5] Between the continued prominence of computer-based learning in many schools due to COVID-19[6] and the ever-increasing amount of time students spend on the Internet and social media,[7] the line of what constitutes activities within the spatial-temporal confines of school is blurry at best. This lack of clarity has created confusion among school officials concerning their ability to discipline students for harmful speech that originates off-campus.[8] Among students, it has led to concerns about when, if ever, they can express themselves freely without fear of punishment from school officials.[9] In the lower courts, this confusion has also led to the emergence of many different approaches governing the discipline of students for their speech—creating a patchwork of fragmented policies across jurisdictions.[10]

The Supreme Court addressed this issue of whether the First Amendment prohibits school officials from regulating speech created by students off-campus for the first time in June 2021 when it decided *Mahanoy Area School District v. B.L.*[11] The decision was announced amidst a time when student Internet usage reached all-time highs, as schools across the

---

4. *See* Alexandra Samur, *The History of Social Media: 29+ Key Moments*, HOOTSUITE (Nov. 22, 2018), https://blog.hootsuite.com/history-social-media/ [https://perma.cc/CY7H-KPNC].

5. *See, e.g.*, Brief for Huntsville, Alabama City Board of Education et al. as Amici Curiae Supporting Petitioner, Mahanoy Area Sch. Dist. v. B.L. *ex rel.* Levy, 141 S. Ct. 2038 (2021) (No. 20-255), 2021 WL 859700, at *10 (listing forty off-campus student speech cases with reported decisions in federal court); *see also* Hewlette-Bullard *ex rel.* J.H-B. v. Pocono Mountain Sch. Dist., 522 F. Supp 3d. 78, 99 (M.D. Pa. 2021); McLaughlin v. Bd. of Regents of Univ. of Okla., 566 F. Supp. 3d 1204, 1213-14 (W.D. Okla. 2021), *appeal docketed*, No. 21-6142 (10th Cir. Oct. 28, 2021); Cheadle *ex rel.* N.C. v. N. Platte R-1 Sch. Dist., 555 F. Supp. 3d 726, 733 (W.D. Mo. 2021), *appeal dismissed*, No. 21-2963, 2021 WL 7186863 (8th Cir. Nov. 2, 2021); McClelland v. Katy Indep. Sch. Dist., No. 4:21-CV-00520, 2021 WL 5055053, at *8-9 (S.D. Tex. Nov. 1, 2021), *appeal docketed*, No. 21-20625 (5th Cir. Nov. 30, 2021; Longoria Next Friend of M.L. v. San Benito Indep. Consol. Sch. Dist., 942 F.3d 258, 264 (5th Cir. 2019); Yeasin v. Durham, 719 F. App'x 844, 849 (10th Cir. 2018).

6. *See* Perry Stein, *Enrollment in Virtual Schools Is Exploding. Will Students Stay Long Term?*, WASH. POST (Feb. 19, 2022, 12:00 PM), https://www.washingtonpost.com/education/2022/02/19/virtual-school-enrollment-increase/ [https://perma.cc/4Z6E-ABJX].

7. *See, e.g.*, Monica Anderson & Jingjing Jiang, *Teens, Social Media and Technology 2018*, PEW RSCH. CTR. 8 (May 31, 2018), https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2018/05/PI_2018.05.31_TeensTech_FINAL.pdf [https://perma.cc/HUJ7-8FZ9].

8 *See, e.g.,* Mahanoy Area Sch. Dist. v. B.L. *ex rel.* Levy*,* 141 S. Ct. 2038, 2063 (2021) (Thomas, J., dissenting).

9. *See* Maureen Downey, *Opinion: Public Schools Can Still Wrongly Punish Off-Campus Student Speech*, ATLANTA J.-CONST.: GET SCHOOLED BLOG (June 28, 2021), https://www.ajc.com/education/get-schooled-blog/opinion-public-schools-can-still-wrongly-punish-off-campus-student-speech/YDJLPRHZPJD4PAXPCJPGOJHCWE/ [https://perma.cc/C6XS-Z6FX].

10. *See, e.g.*, CATHERINE J. ROSS, LESSONS IN CENSORSHIP: HOW SCHOOLS AND COURTS SUBVERT STUDENTS' FIRST AMENDMENT RIGHTS 224-25 (2015) (explaining the approaches taken by each of the circuit courts).

11. 141 S. Ct. at 2045.

country were forced to switch from in-person to online learning to stop the spread of COVID-19.[12] This period was further marked by growing concerns about the adverse effects of social media on youth mental health, as researchers and academics reported connections between increased social media usage among teenagers and elevated rates of anxiety, depression, and body image issues.[13] Due to these circumstances, many had high hopes that the Supreme Court would end this uncertainty surrounding schools' authority to discipline students for their off-campus speech and provide clear guidance for schools and lower courts to rely upon.[14] However, in *Mahanoy*, the Court did anything but—merely providing a highly particularized decision that left for "future cases to decide where, when, and how" schools' regulation of off-campus student speech may violate the First Amendment.[15]

In light of the Internet dramatically expanding the reach of students' speech, the *Mahanoy* opinion's vague description of schools having a "somewhat less[er]" authority to regulate off-campus speech must be clarified to provide school administrators and lower courts with a workable standard for determining what actions are appropriate in the future.[16] Moreover, because of the latitude given to the lower courts to define what these vague standards mean*, Mahanoy* essentially empowers district court judges to give effect to their policy preferences on this issue, creating varied understandings of the scope of students' free speech rights across the country.[17] Because of these problems, this Note proposes that the Supreme Court abandon its current approach of considering the location from which student speech originates. Instead, it argues that the Court should adopt a multi-step sequential evaluation process, modeled mainly after the five-step sequential evaluation process used by the Social Security Administration for disability determinations.[18] This proposed test would provide for greater efficiency, fairness, and predictability among the lower courts.[19] Under this test, students

---

12. *See* Colleen McClain et al., *The Internet and the Pandemic*, PEW RSCH. CTR. 4 (Sept. 1, 2021), https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2021/09/PI_2021.09.01_COVID-19-and-Tech_FINAL.pdf [https://perma.cc/HEA4-9GWW].

13. *See, e.g.*, Deepa Seetharaman, *Senators Seek Answers from Facebook After WSJ Report on Instagram's Impact on Young Users*, WALL ST. J. (Sept. 14, 2021, 8:11 PM), https://www.wsj.com/articles/senators-seek-answers-from-facebook-after-wsj-report-on-instagrams-impact-on-young-users-11631664695 [https://perma.cc/9BZG-RQXN].

14. *See, e.g.*, Frank D. Lomonte, *The Future of Student Free Speech Comes Down to a Foul-Mouthed Cheerleader*, SLATE (Mar. 29, 2021), https://slate.com/technology/2021/03/mahanoy-area-school-district-supreme-court-snapchat-cheerleader.html [https://perma.cc/HP95-5LJX]; *see also* Josh Blackman, *The Incomprehensibility of* Mahanoy Area School District v. B.L., REASON: VOLOKH CONSPIRACY (June 25, 2021, 9:00 AM), https://reason.com/volokh/2021/06/25/the-incomprehensibility-of-mahanoy-area-school-district-v-b-l/ [https://perma.cc/Y8J6-DRXP].

15. 141 S. Ct. at 2046.

16. *Id.* at 2059 (Thomas, J., dissenting).

17. *See* Planned Parenthood of Se. Pa. v. Casey, 505 U.S. 833, 992 (Scalia, J., concurring in part and dissenting in part) (discussing how a lack of concrete guidance on abortion has created a fractured legal regime based upon jurists' individual policy preferences).

18. *See* 20 C.F.R. § 404.1520 (2020).

19. *See* ROBERT L. GLICKSMAN & RICHARD E. LEVY, ADMINISTRATIVE LAW: AGENCY ACTION IN LEGAL CONTEXT 697 (3d ed. 2020).

bear the burden of demonstrating that the speech for which they were disciplined did not have a "sufficient nexus" to the school;[20] or, if there was a nexus, that it did not fall within the categories of speech the Court has deemed to be within the purview of schools to regulate. If the student successfully meets this burden, the burden of proof shifts to the school to show that the challenged speech posed a "reasonably foreseeable risk" of "material disruption" to the school's pedagogical interests.[21]

Before delving into the proposed test, this Note will first provide a brief overview of the First Amendment, Supreme Court precedent governing student speech, the emergence of the Internet and social media, the state of student social media usage, and the lower courts' approaches to regulating off-campus student speech in the Internet era. Next, it will elaborate upon why the Supreme Court's current approach for adjudicating student speech cases is inadequate in terms of providing guidance to students about the scope of their speech rights. This will demonstrate the need for a clarified test to guide school administrators and the lower courts' decision-making processes. This section will further outline the proposed test for evaluating the breadth of schools' authority to regulate off-campus student speech. Finally, this Note will conclude with closing thoughts on the need for the Court to replace the indeterminate guidelines it provided in *Mahanoy* with a more workable test to govern schools' disciplinary authority over off-campus student speech.

II. BACKGROUND

A. The First Amendment and the Right to Free Speech

The First Amendment of the United States Constitution provides that "Congress shall make no law . . . abridging the freedom of speech . . ." [22] As the First Amendment contains no definition of what constitutes "the freedom of speech," our understanding of the scope of this freedom comes from Supreme Court opinions.[23] In this regard, while the language of the First Amendment only explicitly bans *Congress* from taking actions that may chill citizens' speech, the Court has interpreted the free speech rights it confers to be "fundamental personal rights and 'liberties' protected by the Due Process Clause of the Fourteenth Amendment from infringement by the States."[24] Stemming from this recognition of the freedom of speech as a "fundamental right," the Court has understood the right broadly, placing an express prohibition on the government's ability to place constraints on speech because of "its message[,] . . . ideas[,] . . . subject matter, or . . . content."[25] More specifically, it has interpreted the freedom to encompass the freedoms of

20. Kowalski v. Berkeley Cnty. Schs., 652 F.3d 565, 577 (4th Cir. 2011).

21. Wisniewski v. Bd. of Educ. of Weedsport Cent. Sch. Dist., 494 F.3d 34, 38-39 (2d Cir. 2007) (quoting Tinker v. Des Moines Indep. Cmty. Sch. Dist., 393 U.S. 503, 513 (1969)).

22. U.S. CONST. amend. I.

23. *See* GREGORY E. MAGGS & PETER J. SMITH, CONSTITUTIONAL LAW: A CONTEMPORARY APPROACH 1000 (5th ed. 2021).

24. Gitlow v. New York, 268 U.S. 652, 666 (1925).

25. Police Dep't of Chi. v. Mosley, 408 U.S. 92, 95 (1972).

inquiry and thought,[26] including the "right to speak freely and . . . to refrain from speaking;"[27] the rights to utter, print, and read;[28] the right to distribute and receive literature;[29] and the "right to attempt to persuade others to change their views," even when the speaker's message may offend their audience.[30]

However, despite its robust protections of the freedom of speech, the Supreme Court has not interpreted the First Amendment to confer an absolute right,[31] instead identifying "narrowly limited" classes of unprotected speech.[32] As of April 2022, the Supreme Court has recognized eight categories of speech unprotected by the First Amendment—(1) obscenity, (2) defamation, (3) fraud, (4) incitement, (5) fighting words, (6) true threats, (7) speech integral to criminal conduct, and (8) child pornography.[33] While the Court has acknowledged that there may be additional categories of unprotected speech,[34] it has indicated a "reluctan[ce] to mark off new categories of speech for diminished constitutional protection."[35] Notwithstanding this reluctance, the Court has qualified the breadth of free speech rights as it pertains to children and minors due to their being subject to the control of their parents and guardians until reaching the age of majority.[36]

B. Tracing the Extension of Constitutional Rights to Children and Students

For much of early American history, the law failed to recognize children as having rights apart from their parents or the state, embracing the notion that children were entitled only to be heard through their parents or elders.[37] It was not until the 1960s that the Supreme Court began to explicitly reference children as being holders of their own constitutional rights—declaring that "neither the Fourteenth Amendment nor the Bill of Rights is for adults alone."[38] In the decades that followed, this understanding continued to prevail, with Justice Blackmun further proclaiming that "[c]onstitutional rights do not mature and magically come into being only when one attains the

26. Griswold v. Connecticut, 381 U.S. 479, 482 (1965).
27. Janus v. Am. Fed'n of State, Cnty. & Mun. Emps., Council 31, 138 S. Ct. 2448, 2463 (2018) (quoting Wooley v. Maynard, 430 U.S. 705, 714 (1977)).
28. *Griswold*, 381 U.S. at 482.
29. *See* Martin v. City of Struthers, 319 U.S. 141, 143 (1943).
30. Hill v. Colorado, 530 U.S. 703, 716 (2000).
31. *See Gitlow*, 269 U.S. at 666.
32. Chaplinsky v. New Hampshire, 315 U.S. 568, 571-72 (1942).
33. *See* VICTORIA L. KILLION, CONG. RSCH. SERV., IF11072, THE FIRST AMENDMENT: CATEGORIES OF SPEECH (2019).
34. *See* United States v. Stevens, 559 U.S. 460, 472 (2010).
35. Nat'l Inst. of Fam. & Life Advocs. v. Becerra, 138 S. Ct. 2361, 2371-72 (2018).
36. *See, e.g.*, Vernonia Sch. Dist. 47J v. Acton, 515 U.S. 646, 654-55 (1995).
37. *See* DOUGLAS E. ABRAMS ET AL., CHILDREN AND THE LAW IN A NUTSHELL 9-17 (7th ed. 2021).
38. Laurence D. Houlgate, *Three Concepts of Children's Constitutional Rights: Reflections on the Enjoyment Theory*, 2 U. PA. J. CONST. L. 77 (1999) (quoting *In re* Gault, 387 U.S. 1, 13 (1967)).

state-defined age of majority. Minors, as well as adults, are protected by the Constitution and possess constitutional rights."[39]

Despite this recognition of children as possessors of constitutional rights, the Court has clarified that their enjoyment of such rights is not the same as adults.[40] Observing that children are not capable of taking care of themselves[41] due to their "peculiar vulnerabilit[ies]" and inability to make mature and informed decisions, the Court has reasoned it would be inappropriate to recognize constitutional protections afforded to adults as robustly for children.[42]

In the context of school, the Court has similarly applied this understanding of children possessing rights of a "lesser magnitude" than adults[43] to justify school officials' tutelary control over their students.[44] Based on its view of schools having the duty to instill "habits and manners of civility"[45] and teach cultural values necessary for students' development into adults,[46] the Court has long utilized the English common law doctrine of *in loco parentis*[47] to provide school officials with the authority to maintain order within their schools.[48] In applying this doctrine, the Court has acknowledged that "while children assuredly do not 'shed their constitutional rights . . . at the schoolhouse gate,' the nature of those rights is what is appropriate for children in school."[49] Through this understanding of schools' standing *in loco parentis* over their students, the Court has further justified granting school officials "First Amendment leeway" to discipline behaviors that occur under their supervision—deeming deviations from traditional First Amendment doctrine to be permissible when necessary to protect the "special characteristics of the school environment."[50] Using this reasoning, it has permitted school officials to prohibit the use of "vulgar and offensive terms in public discourse,"[51] and speech that is "reasonably viewed as promoting illegal drug use" in school or at school events.[52] In addition, it has further justified school officials to "censor school-sponsored publications . . .

39. Planned Parenthood of Central Missouri v. Danforth, 428 U.S. 52, 74 (1976).

40. *See, e.g.*, Prince v. Massachusetts, 321 U.S. 158, 168 (1944); *Developments in the Law: The Constitution and the Family*, 93 HARV. L. REV. 1156, 1358 (1980) [hereinafter *Developments*].

41. *See* Schall v. Martin, 467 U.S. 253, 265 (1984).

42. Bellotti v. Baird, 443 U.S. 622, 634 (1979).

43. *Developments*, *supra* note 40, at 1358.

44. *See* Vernonia Sch. Dist. 47J v. Acton, 515 U.S. 646, 655 (1995).

45. Bethel Sch. Dist. No. 403 v. Fraser, 478 U.S. 675, 681 (1986).

46. *See* Brown v. Bd. of Educ., 347 U.S. 483, 493 (1954).

47. *See In Loco Parentis*, BLACK'S LAW DICTIONARY (11th ed. 2019). Translated from Latin, "*in loco parentis*" means "in the place of a parent." *Id.*

48. *See e.g.*, *Acton*, 515 U.S. at 655-56; *see also* Morse v. Frederick, 551 U.S. 393, 413 (2007) (Thomas, J., concurring).

49. *Acton*, 515 U.S. at 655-56 (quoting Tinker v. Des Moines Indep. Cmty. Sch. Dist., 393 U.S. 503, 516 (1969)).

50. Mahanoy Area Sch. Dist. v. B.L. *ex rel.* Levy, 141 S. Ct. 2038, 2044-46 (2021) (citing Hazelwood Sch. Dist. v. Kuhlmeier, 484 U.S. 260, 266 (1988)).

51. *Fraser*, 478 U.S. at 676.

52. *Morse,* 551 U.S. at 403.

reasonably related to legitimate pedagogical concerns"[53] or "other expressive activities . . . members of the public might reasonably perceive to bear the imprimatur of the school" as within the scope of schools' disciplinary authority.[54]

Concerning this additional latitude afforded to schools to maintain discipline and order, the Court has repeatedly justified this greater degree of control over student expression as necessary to protect the "special characteristics of the school environment."[55] Scholars and commentators have interpreted these characteristics to include: the age and maturity of schoolchildren; the fact that, for many students, school attendance is made compulsory by law; how schools serve the sometimes-competing interests of the parents, children, and the state; the heightened safety considerations required of school administrators; the expectation of public accountability; and the need to promote educational goals.[56]

In *Tinker v. Des Moines Independent Community School District,* commonly regarded as the foundational case for students' rights, the Supreme Court famously declared, "It can hardly be argued that either students or teachers shed their constitutional freedom of speech or expression at the schoolhouse gate."[57] In *Tinker*, students suspended for wearing black armbands to school to protest the Vietnam War sued their school district, arguing that the suspension violated their First Amendment free speech rights.[58] Addressing the school's authority to regulate the students' speech, the Court held that the First Amendment barred school officials from censoring student speech on or off campus unless such speech "might reasonably have led school authorities to forecast substantial disruption of or material interference, with school activities" or a showing that a disturbance on school premises actually occurred.[59] The Court in *Tinker* further explained that these protections were not limited to the classroom, but extended to all school facilities and established that schools could not discipline students for simply expressing opposing viewpoints that create discomfort.[60]

While *Tinker* is lauded by many as a decision protective of student speech rights—due to its enumeration of them in the first place—many fail to recognize the limitations on student speech it also created.[61] Specifically, in its phrasing of the oft-cited substantial disruption test, the Court constrained the speech rights of students.[62] By stating that "conduct by the student in class or out of it" was not subject to First Amendment protection, the Court significantly expanded the realm of behaviors within schools' disciplinary

53. *Acton,* 515 U.S. at 655-56 (citing *Kuhlmeier*, 484 U.S. at 273).
54. *Kuhlmeier*, 484 U.S. at 271.
55. *Mahanoy*, 141 S. Ct. at 2044-46 (citing *Kuhlmeier*, 484 U.S. at 266).
56. *See* Bryan R. Warnick, *Student Speech Rights and the Special Characteristics of the School Environment*, 38 EDUC. RESEARCHER 200, 201 (2009).
57. 393 U.S. 503, 506 (1969).
58. *See id.* at 504.
59. *Id.* at 514.
60. *See id.* at 509, 513.
61. *See* Mary-Rose Papandrea, *The Great Unfulfilled Promise of* Tinker, 105 VA. L. REV. 159, 159-60 (2019).
62. *See id*.

authority.[63] In addition, by permitting school officials to act when they "might reasonably . . . forecast substantial disruption of . . . school activities," the Court left much discretion to schools to determine what expressive activities created a sufficient level of foreseeable disruption, as opposed to only authorizing discipline for harm that had occurred.[64] These limitations have become even more pronounced in the last few decades, as students' Internet and social media usage has facilitated more opportunities for off-campus student speech than ever before.[65]

C. The Internet and Contemporary Forms of Student Speech

1. The Emergence of the Internet and Social Media

Historians and academics alike regard October 29, 1969, the day the first message was delivered through an interconnected computer network, as the day the modern Internet was born.[66] The Internet has developed and grown immensely in the five decades since, revolutionizing how we live and communicate.[67] Today's Internet can hardly be cabined to being merely an "electronic communications network" that connects people around the world, as defined in Merriam-Webster's online dictionary;[68] rather, today almost anything from watching movies to banking to even ordering groceries can be done online.[69]

With the emergence of the Internet, so too emerged many new forms of media, including e-mail, instant messaging, and social media.[70] These new media forms have further contributed to the revolution spawned by the Internet by providing more accessible and faster ways to communicate and share information.[71] Through social media—defined as "websites and other

63. *Tinker*, 393 U.S. at 513*; see also* Papandrea, *supra* note 61, at 159-60.

64. *See* Ben Lee, *What* Tinker *Got Wrong,* FOUND. FOR INDIVIDUAL RTS. EDUC. (Sept. 28, 2018), https://www.thefire.org/what-tinker-got-wrong/ [https://perma.cc/62CE-26HU]; *see also* Papandrea, *supra* note 61, at 170-71.

65. *See* Beth A. Narrow & Sommer Ingram Dean, *The Law of Students' Rights to Online Speech: The Impact of Students' Ability to Openly Discuss Public Issues*, HUM. RTS. MAG., Jan. 2022, at 17.

66. *See* Matt Blitz, *What Will the Future of the Internet Look Like?*, POPULAR MECHS. (Sept. 30, 2021), https://www.popularmechanics.com/technology/infrastructure/a29666802/future-of-the-internet/ [https://perma.cc/4YFR-FCG2].

67. *See id.*

68. *Internet*, MERRIAM-WEBSTER.COM, https://www.merriam-webster.com/dictionary/Internet [https://perma.cc/KZ4X-S5HN] (last visited Nov. 12, 2021).

69. *See* Ella Koeze & Nathaniel Popper, *The Virus Changed the Way We Internet*, N.Y. TIMES (Apr. 7, 2020), https://www.nytimes.com/interactive/2020/04/07/technology/coronavirus-internet-use.html [https://perma.cc/9CT2-VCCX].

70. *See* Michael Aaron Dennis et al., *Internet*, ENCYC. BRITANNICA, https://www.britannica.com/technology/Internet [https://perma.cc/3EZH-TMNG] (last visited Jan. 25, 2022).

71. *See* Sol Rogers, *The Role of Technology in the Evolution of Communication*, FORBES (Oct. 15, 2019, 8:57 AM), https://www.forbes.com/sites/solrogers/2019/10/15/the-role-of-technology-in-the-evolution-of-communication/ [https://perma.cc/U8GD-UALW].

online means of communication that large groups of people use to share information and develop social and professional contacts"—users can share photos and videos and communicate with their friends and family from wherever they have cellphone service or Internet connection.[72]

In the last few years, the Supreme Court has begun to acknowledge these increasingly prominent forms of media, even formally recognizing sentiments communicated on the Internet and social media as free speech activities protected by the First Amendment.[73] In *Packingham v. North Carolina*, Justice Kennedy remarked about how social media enables anyone with working Internet "to become a town crier with a voice that resonates farther than it could from any soapbox."[74] With recent surveys indicating that over 72% of Americans report using at least one social media platform—the impact and reach of citizen's speech is only likely to continue growing.[75]

2. Student Social Media Use Today

Reflective of the trends in Internet use among the population at large, the Internet and social media play even more significant roles in the lives of American youth.[76] Among teenagers, YouTube,[77] Instagram,[78] Snapchat,[79] Facebook,[80] and Twitter[81] are the most popular online platforms.[82] Most popular social media companies, including YouTube, Instagram, Facebook,

---

72. *Social Media,* DICTIONARY.COM, https://www.dictionary.com/browse/social-media [https://perma.cc/44RR-XWAV] (last visited Jan. 25, 2022).

73. *See, e.g.*, Packingham v. North Carolina, 137 S. Ct. 1730, 1731 (2017).

74. *Id.*

75. *See Social Media Fact Sheet*, PEW RSCH. CTR. (Apr. 7, 2021), https://www.pewresearch.org/internet/fact-sheet/social-media/ [https://perma.cc/X7DU-CM7J]; *Packingham*, 137 S. Ct. at 1731.

76. *See* Anderson & Jiang, *supra* note 7; VICTORIA RIDEOUT & MICHAEL B. ROBB, THE COMMON SENSE CENSUS: MEDIA USE BY TWEENS AND TEENS, 3 (Jenny Pritchett ed., 2019), https://www.commonsensemedia.org/sites/default/files/research/report/2019-census-8-to-18-full-report-updated.pdf [https://perma.cc/UXL7-7J8U].

77. YouTube is a website where users can watch or share videos. *YouTube*, DICTIONARY.COM, https://www.dictionary.com/browse/youtube [https://perma.cc/YDV5-GE7K] (last visited Oct. 7, 2022).

78. Instagram is an application on which users can share photos and videos with their friends. Elise Moreau, *What Is Instagram and Why Should You Be Using It?*, LIFEWIRE (Sept. 12, 2021), https://www.lifewire.com/what-is-instagram-3486316 [https://perma.cc/C6E6-2EBK].

79. *See* Mahanoy Area Sch. Dist. v. B.L. *ex rel.* Levy*,* 141 S. Ct. 2038, 2040 (2021) (describing Snapchat as "a social media application for smartphones that allows users to share temporary images with selected friends").

80. Facebook is a website where users can connect with friends, post comments, share photographs, news clips, linked to other websites, and other content either to either select people, groups of friends, or the public at large. Daniel Nations, *What Is Facebook? Learn Why So Many People Can't Stay Away from Facebook*, LIFEWIRE (Sept. 19, 2021), https://www.lifewire.com/what-is-facebook-3486391 [https://perma.cc/TET9-RY3S].

81. Twitter is an online news and social networking application where people communicate in messages limited to 280 characters. Paul Gil, *What Is Twitter & How Does It Work?*, LIFEWIRE (Aug. 30, 2021), https://www.lifewire.com/what-exactly-is-twitter-2483331 [https://perma.cc/B5N3-LVMM].

82. *See* Anderson & Jiang, *supra* note 7, at 2.

and Twitter, require users to be thirteen or older to make an account because of the Children's Online Privacy Protection Act's prohibition of website operators from collecting information from children under thirteen.[83] However, children ages eight to twelve have been easily able to get around these age restrictions—with reports indicating 76% of children in this age group use YouTube[84] and as much as 50% of children ages eleven and twelve have social media profiles.[85]

Looking at social media usage patterns of school-aged children more broadly, surveys of American adolescents aged eight to eighteen years old indicate that exclusive of time spent on digital devices for school and homework, children aged eight to twelve spend approximately four hours and forty-four minutes on screen media, and teenagers aged thirteen to eighteen spend about seven hours and twenty-two minutes on screen media.[86] Further, 95% of American teenagers report having access to a smartphone, and 89% of teenagers claim to use the Internet at least "several times a day." [87] With students' almost constant use of their phones, the Internet, and social media, most students "are engaging in enormous amounts of off-campus speech."[88]

Because of the far greater reach and speed at which Internet-generated speech can be received, speech that a student posts or sends while off campus is regularly received by fellow students on-campus.[89] Given the frequency with which students have brought challenges against their schools for discipline related to their Internet speech, the lower courts have had to grapple with the limited guidance provided by the Court to determine how to best adjudicate these issues in their jurisdictions.[90]

---

83. Children's Online Privacy Protection Act, 15 U.S.C. § 6502 (2012).

84. RIDEOUT & ROBB, *supra* note 76, at 34.

85. *See, e.g.*, *Under-Age Social Media Use 'On the Rise', Says Ofcom*, BBC NEWS (Nov. 29, 2017), https://www.bbc.com/news/technology-42153694 [https://perma.cc/M5QJ-LDCJ]; Eleanor Harding, *Six in Ten Parents Say They Would Let Their Children Lie About Their Age Online to Access Social Media Sites*, DAILY MAIL (Jan. 24, 2017, 2:48 AM), https://www.dailymail.co.uk/news/article-4150204/Many-parents-let-children-lie-age-online.html [https://perma.cc/TVC7-5ZE9].

86. RIDEOUT & ROBB, *supra* note 76, at 3. Included in its term "screen media," the article references several activities including watching tv and videos, playing video games, using social media, listening to music, reading, writing, video chatting, browsing, and creating content. *See id*. at 6.

87. Anderson & Jiang, *supra* note 7, at 2, 8.

88. Brief for Independent Women's Law Center as Amicus Curiae Supporting Respondents, Mahanoy Area Sch. Dist. v. B.L. *ex rel.* Levy, 141 S. Ct. 2038 (2021) (No. 20-255), 2021 WL 1255349, at *16-17.

89. *Mahanoy*, 141 S. Ct. at 2062-63 (Thomas, J., dissenting).

90. *See* Brief of Huntsville, Alabama City Board of Education et al., *supra* note 5, at *11.

D. Student Speech in the Internet World: Approaches to Regulating Off-Campus Speech

1. Pre-Mahanoy Circuit Court Approaches

As the Internet transformed modern methods of communication, greatly expanding the reach of students' expressive activity,[91] the lower courts had nothing more than the broad statement from *Tinker* that "conduct by the student in class or out of it" could be punished by school officials if it created or threatened a sufficient risk of substantial disruption to the school to guide them.[92] With such indeterminate instructions, the lower courts were left to their own devices to determine what constituted on- versus off-campus speech, and what behaviors were sufficient to satisfy this "substantial disruption" standard.[93] From this uncertainty, three predominant approaches emerged among the circuit courts—the reasonable foreseeability test, the sufficient nexus test, and an approach entirely rejecting the applicability of *Tinker* to off-campus speech.[94]

a. The Reasonable Foreseeability Test

The reasonable foreseeability test has been the most popular standard for applying *Tinker* to off-campus speech among the circuit courts, with the Second, Eighth, and Eleventh Circuits applying it to guide their determinations.[95] Under this test, schools may regulate students' off-campus speech when it is "reasonably foreseeable" that the student's communication will "substantially disrupt the work and discipline of the school" environment.[96]

In *Wisniewski v. Board of Education of the Weedsport Central School District*, the case credited with creating this test, the Second Circuit was faced with determining whether a student's instant messages sent from his home computer were within the school's authority to regulate.[97] The messages at issue included a picture of a pistol firing a bullet at a person's head with the caption "Kill Mr. VanderMolen."[98] The Second Circuit dismissed the

91. *See* Rogers, *supra* note 71.

92. Tinker v. Des Moines Indep. Cmty. Sch. Dist., 393 U.S. 503, 513 (1969).

93. Daniel Marcus-Toll, Note, Tinker *Gone Viral: Diverging Threshold Tests for Analyzing School Regulation of Off-Campus Digital Student Speech*, 82 FORDHAM L. REV 3395, 3409 (2014); Lee, *supra* note 64.

94. *See, e.g.*, ROSS, *supra* note 10, at 224-25; Nicolas Burnosky, Comment, *2-4-6-8 Who Do We Appreciate? The Third Circuit Scores a Touchdown for Student-Athlete Free Speech Rights*, 28 JEFFREY S. MOORAD SPORTS L.J. 369, 380 (2021). While the concepts guiding these three tests are recognized as the predominant circuit court tests, they are not uniformly titled as they are in this Note.

95. *See* Meghan K. Lawrence, Note, Tinker *Stays Home: Student Freedom of Expression in Virtual Learning Platforms*, 101 B.U. L. REV. 2249, 2265-66 (2022).

96. Wisniewski v. Bd. of Educ. of Weedsport Cent. Sch. Dist., 494 F.3d 34, 38-39 (2d Cir. 2007) (quoting Tinker v. Des Moines Indep. Cmty. Sch. Dist., 393 U.S. 503, 513 (1969)).

97. *See id.*

98. *Id.* at 36.

student's First Amendment claim against his school officials, finding the messages to be unprotected based on the substantial disruption framework provided in *Tinker*.[99] It opined that no reasonable jury could conclude that it was unforeseeable that the student's messages would come to the attention of school officials and create a substantial disruption to the work and discipline of the school.[100]

b. The Sufficient Nexus Test

Under the sufficient nexus test, which was introduced by the Fourth Circuit, schools may discipline students for off-campus speech when there is a close connection between the speech and the school's pedagogical interests.[101] In *Kowalski v. Berkeley County Schools*, the case to which this test is attributed, the Fourth Circuit applied this substantial nexus test to determine whether a student's suspension for her off-campus social media activity that targeted and referred to one of her classmates as being a "slut" and having herpes violated her First Amendment rights.[102] Finding that the student's free speech rights were not violated, the Fourth Circuit invoked *Tinker* and reasoned that even though the ability of schools to regulate students' off-campus speech is not unlimited, here, the nexus of the student's social media activity to the school and the subsequent interference it caused within the school were sufficient to justify its disciplinary action.[103] The Fourth Circuit concluded the opinion with a declaration that where student "speech has a sufficient nexus with the school, the Constitution is not written to hinder school administrators' good faith efforts to address the problem."[104]

The Ninth Circuit has also applied the sufficient nexus test, but under its understanding of the test, a totality of the circumstances inquiry is required to determine if the student's speech is closely connected to the school.[105]

c. The "Tinker is Inapplicable to Off-Campus Speech" Approach

Under this approach, used primarily by the Third Circuit, judges reject the idea that *Tinker* authorized schools to regulate off-campus speech.[106] This reading of *Tinker* and ultimate refusal to recognize schools' authority to discipline students for off-campus speech stems from a fear that doing so would allow "the state, in the guise of school authorities to reach into a child's

99. *See id.* at 35.

100. *See id.* at 39-40.

101. *See* Marcus-Toll, *supra* note 93, at 3420; THOMAS A. YOUNG, LEGAL RIGHTS OF CHILDREN § 17:3 (3d ed. 2021).

102. 652 F.3d 565, 567-69 (4th Cir. 2011).

103. *See id.* at 572-73.

104. *Id.* at 577.

105. *See* McNeil v. Sherwood Sch. Dist. 88J, 918 F.3d 700, 707-08, 712 (9th Cir. 2019) (internal citations omitted).

106. *See* ROSS, *supra* note 10, at 225.

home and control" their actions.[107] Judges advocating for this approach have also argued that students' off-campus speech should receive the same protections as adults, reasoning that the "special characteristics of the school environment" justifying lesser protection for student speech in schools are absent outside the "schoolhouse gate."[108]

In *B.L. v. Mahanoy Area School District*, the appellate level case that preceded *Mahanoy*, the Third Circuit held that *Tinker* does not apply to off-campus speech—defining off-campus speech to include any speech made outside of school-owned, -operated, or -supervised channels.[109] It reasoned that doing so would offer greater clarity to students, as it would be much easier for them to determine whether their speech occurred in a school-operated setting than if the speech had some indeterminate "nexus" to the school.[110] However, this approach was explicitly rejected by the majority in *Mahanoy*, who reasoned that certain speech that originates off-campus may still constitute important regulatory interests for the school, such as severe bullying or harassment.[111]

2. Mahanoy Area School District v. B.L.

In *Mahanoy Area School District v. B.L.*, the Supreme Court addressed the question of whether a school could punish a student for speech made while off-campus for the first time.[112] The case centered around a high school student's claim that her school district violated her free speech rights by suspending her from the junior varsity cheerleading team following her sending two Snapchat messages while off-campus one weekend.[113] The messages at issue had been posted to the student's Snapchat story after she learned she was not selected for either her school's varsity cheerleading team or her desired position on her school's softball team.[114] One of the messages contained text indicating the student's anger about not making the varsity cheerleading team, while the other had an image of her and a friend accompanied by the caption "f**k school f**k softball f**k cheer f**k everything."[115] After the images spread, the coaches of the junior varsity cheerleading team, in consultation with the school, suspended the student from the team for the upcoming school year.[116] She and her parents subsequently filed suit in the district court.[117]

---

107. *Id.* (quoting Layshock *ex rel.* Layshock v. Hermitage Sch. Dist., 650 F.3d 205, 216 (3d Cir. 2011)).

108. J.S. *ex rel.* Snyder v. Blue Mountain Sch. Dist., 650 F.3d 915, 936 (3d Cir. 2011) (Smith, J., concurring).

109. *See* B.L. *ex rel.* Levy v. Mahanoy Area Sch. Dist., 964 F.3d 170, 189 (3d Cir. 2020), *aff'd on other grounds*, 141 S. Ct. 2038 (2021).

110. *Id.* at 189-90.

111. Mahanoy Area Sch. Dist. v. B.L. *ex rel.* Levy, 141 S. Ct. 2038, 2045 (2021).

112. *See id.* at 2044.

113. *See id.* at 2043.

114. *See id.*

115. *Id.*

116. *See id.*

117. *Mahanoy*, 141 S. Ct. at 2043.

Writing for an 8-1 majority, Justice Breyer held that the suspension violated the student's First Amendment free speech rights, reasoning that the school was unable to demonstrate that there was either a reasonable threat or occurrence of a "substantial disruption" because of the offending Snapchat messages.[118] He further rejected the school's purported interest in teaching civility and good manners, deeming it an insufficient interest to overcome the student's right to free speech.[119]

Aside from reaffirming the applicability of the *Tinker* "substantial disruption" test to off-campus speech, the majority opinion provided little additional guidance as to what kind of off-campus speech would constitute a sufficient disruption. The opinion merely mentioned three features of off-campus speech that "diminish the unique educational characteristics that might call for special First Amendment leeway."[120] Specifically, Breyer referenced three attributes: (1) the fact that the doctrine *in loco parentis* is generally inapplicable to off-campus student speech; (2) the concern that imposing restrictions on students' off-campus speech would subject students to speech restrictions twenty-four hours a day, having a serious chilling effect; and (3) the observation that schools have an important duty to protect students who espouse unpopular ideas as a means to promote the continued preservation of a well-informed, democratic society.[121] Yet, like with the rest of the considerations he mentions in the opinion, Justice Breyer declined to assign determinative values to these characteristics or even to define off-campus speech; leaving the matter for future cases to decide.[122] Thus, in place of formal guidance, he offered a list of off-campus student conduct illustrative of what might be permissible for schools to regulate—including severe bullying, threats to fellow students or teachers, and breaches of school security devices.[123] Because *Mahanoy* provides little more than these broad declarations of principles, lower courts are left with no clear standards to guide future cases.[124]

3. Confusion in the Lower Courts Post-Mahanoy

Stemming from the indeterminate guidance provided by *Mahanoy*, lower courts addressing similar issues in its wake continue to be inconsistent in determining when schools' regulation of off-campus student speech is constitutional.[125] At least one district court in the Tenth Circuit has read *Mahanoy*'s protection of students' off-campus speech broadly, interpreting

118. *See id.* at 2047-48.
119. *See id.*
120. *Id.* at 2045-46.
121. *See id.*
122. *See id.*
123. *Mahanoy*, 141 S. Ct. at 2045-46
124. *Id.* at 2046; Downey, *supra* note 9.
125. *See, e.g.*, McLaughlin v. Bd. of Regents of Univ. of Okla., 566 F. Supp. 3d 1204, 1213-14 (W.D. Okla. 2021), *appeal docketed*, No. 21-6142 (10th Cir. Oct. 28, 2021); Cheadle on behalf of N.C. v. N. Platte R-1 Sch. Dist., 555 F. Supp. 3d 726, 733 (W.D. Mo. 2021), *appeal dismissed*, No. 21-2963, 2021 WL 7186863 (8th Cir. Nov. 2, 2021).

the case to mean that nearly all student posts on social media that originate off-campus are protected speech.[126] In contrast, another district court in the Fourth Circuit has construed *Mahanoy* more narrowly—finding the school's strong interest in deterring alcohol abuse among its students as a sufficient interest to overcome the "substantial disruption" test.[127] Thus, the court enabled the school to discipline a student for Snapchat videos she had sent of herself drinking in her bedroom, reasoning it was one such regulatory interest Justice Breyer had indicated as remaining "significant" off-campus in *Mahanoy*.[128]

Other courts, such as one district court in the Fifth Circuit, have affirmatively called attention to the lack of clarity provided by *Mahanoy.*[129] In a case decided nearly five months after *Mahanoy*, the district court judge refused to even address the merits of a student's First Amendment claims, reasoning that school officials were shielded by qualified immunity for their actions, as there was no "rule that could have put [them] on notice that it would be unconstitutional" to discipline a student for his sending an offensive Snapchat video to another student off-campus after a football game.[130] The judge harped upon *Mahanoy's* failure to establish a clear rule governing school officials' ability to discipline off-campus speech, making reference to the *Mahanoy* majority's reference to circumstances that "may implicate a school's regulatory interests" without giving any specific criteria.[131]

Because of the variability in outcomes in the lower courts, students are left in limbo about when and where they can express themselves freely without fear of repercussion from school officials.[132] This uncertainty, aside from having a chilling effect on student speech, keeps the door wide open for the continued use of the qualified immunity defense by school officials whenever they face challenges for disciplinary actions concerning speech—even those that violate students' constitutional rights.[133] This reality illustrates the pressing need for the Court to issue a clarified test.[134] The forthcoming

126. *See McLaughlin*, 566 F. Supp. 3d at 1213-14 (referencing *Mahanoy* and stating that "[i]f a student's posting via social media of a direct vulgar attack on her school and its coaches is protected speech . . . , it is difficult to see how posting a somewhat ambiguous emoji on a third-party website . . . could be otherwise.").

127. *See Cheadle*, 555 F. Supp. 3d at 732 (finding no free speech violation when a school suspended a student for sending Snapchat videos of herself drinking alcohol from her bedroom to her classmates because the school's interest in deterring middle schoolers from underage drinking was one of the permissible "significant . . . off-campus circumstances" Justice Breyer authorized in *Mahanoy*).

128. *See id.*

129. *See* McClelland v. Katy Indep. Sch. Dist., No. 4:21-CV-00520, 2021 WL 5055053, at *8-9 (S.D. Tex. Nov. 1, 2021), *appeal docketed*, No. 21-20625 (5th Cir. Nov. 30, 2021).

130. *Id.* at *9.

131. *Id.* at *8 (quoting *Mahanoy*, 141 S. Ct. at 2046).

132. *See* Downey, *supra* note 9.

133. *See* David L. Hudson Jr., *Qualified Immunity*, FIRST AMEND. ENCYC., https://www.mtsu.edu/first-amendment/article/1560/qualified-immunity [https://perma.cc/TY67-Q4B3] (last visited Oct. 7, 2022).

134. *See* Downey, *supra* note 9.

section provides a proposed alternative test for the Court to adopt when it is faced with the next off-campus student speech case.

III. ANALYSIS

A. *The Modified Test: A Systematic Inquiry Assessing the Scope of Schools' Authority to Regulate Student Speech*

To mend the troubling reality that students receive differing protection for their speech based on their circuit's interpretation of the *Tinker* substantial disruption test,[135] the Court should fill the gaps left by *Tinker* and *Mahanoy* by articulating a multi-step sequential evaluation process. While this proposal advocates for the Court to abandon consideration of geographic origin in its evaluation of whether actions taken by school officials are permissible, this new approach is not novel. Instead, it is mainly based upon Supreme Court precedent and dicta indicating behaviors explicitly or implicitly regarded as within or outside the regulatory authority of schools.[136]

The proposed test, a four-step inquiry modeled to function like the Social Security Administration's (the "SSA") five-step sequential evaluation procedure for disability determinations, [137] similarly involves following a series of steps in a set order that functions formulaically for all courts. As was the purpose of the SSA's evaluation process, this test aims to promote efficiency, fairness, and uniformity among courts.[138] For example, suppose a court finds that the challenged speech is among the types of speech recognized by the Supreme Court as within the scope of schools' disciplinary authority in Table B's Step 3, *infra*. In that case, the court would end its inquiry and issue an opinion in favor of the school. In contrast, if at Step 3 the court fails to make a definitive determination based upon the grids in Tables A-C, *infra*, it would proceed to the next and final step in the evaluation process to conclude its inquiry.

Embedded in each of the test's steps are behaviors explicitly or implicitly regarded by the Court as being within or outside the regulatory authority of school officials. In addition, the proposed test consolidates elements of the prevailing circuit court tests for regulating off-campus speech to create one all-encompassing inquiry. A more in-depth explanation of how a court would proceed through each step of the proposed test is provided below.

135. *See* Marcus-Toll, *supra* note 93, at 3436-37.
136. *See infra* Tables B-C.
137. 20 C.F.R. § 404.1520 (2020).
138. *See* GLICKSMAN & LEVY, *supra* note 19, at 697.

1. Step 1: Did the Speech Have a "Sufficient Nexus" to the School?

As a threshold question, a court must first ask whether there is a "sufficient nexus"[139] or close connection between the challenged speech and the school's pedagogical interests. To assess if such a nexus exists, it should look at the totality of the circumstances surrounding the student's speech, engaging in an in-depth fact-specific inquiry into the case at hand.[140] Relevant considerations for this analysis should include whether the speech: (1) bears the "imprimatur" of the school or is proffered through some platform with the school's name or logo;[141] (2) was made during a time when the school was responsible for the student;[142] (3) was made while the students were on their way to or from the school;[143] (4) took place on school grounds, property or digital platforms (such as the school's Zoom account);[144] (5) occurred during in-person or remote instruction;[145] (6) occurred during extracurricular activities sponsored or offered by the school;[146] (7) identified the school or targeted a member of the school community with vulgar or abusive language;[147] or (8) involved a failure to follow rules concerning school assignments.[148] While not an exhaustive list, this suggested inquiry consolidates considerations advanced by the Supreme Court, and expanded upon by lower courts, in deciding what constitutes a connection to the school significant enough to warrant punishment.

For a student's speech to have a sufficient nexus to the school, it need not meet all the above-listed considerations. Instead, each of the factors present should be considered cumulatively to assess its relative connection to the school—with a "sufficient nexus" being found where the balance of the scale is titled toward connection to the school. If, after this totality of the circumstances analysis, a court determines the speech has a sufficient connection to the school, it should move on to the next step of the evaluation process. If, however, the speech does not have a sufficient nexus to the school, a court must dismiss the case in favor of the student, as the school cannot regulate speech that is "in no way connected with or affecting the school," for the discipline of such conduct falls within the zone of parental authority.[149]

This step incorporates the "sufficient nexus" test applied by the Fourth and Ninth Circuits.[150] While this test has been subject to criticism for

139. Kowalski v. Berkeley Cnty. Schs., 652 F.3d 565, 577 (4th Cir. 2011).

140. McNeil v. Sherwood Sch. Dist. 88J, 918 F.3d 700, 707 (9th Cir. 2019).

141. Hazelwood Sch. Dist. v. Kuhlmeier, 484 U.S. 260, 273 (1988).

142. Mahanoy Area Sch. Dist. v. B.L. *ex rel.* Levy, 141 S. Ct. 2038, 2054 (2021) (Alito, J., concurring).

143. *Id.*

144. *Id*. at 2045 (majority opinion).

145. *Id.* at 2054 (Alito, J., concurring).

146. *Id.*

147. *Id.* at 2045 (majority opinion).

148. *Manahoy*, 141 S. Ct. at 2045.

149. *Id.* at 2060 (Thomas, J., dissenting) (quoting Lander v. Seaver, 32 Vt. 114, 120 (1859)).

150. *Kowalski*, 652 F. 3d at 577; *see also McNeil*, 918 F.3d at 707-08, 712.

affording little clarity to students on what speech could subject them to punishment,[151] the addition of the eight suggested factors to guide a court's determination offers students greater guidance of what speech may subject them to punishment.[152] In addition, by requiring such an in-depth case-by-case inquiry, this step seeks to add a layer of protection for students, ensuring they can only be disciplined for speech that is within the school's regulatory purview.

2. Step 2: Did the Speech Implicate a Matter of Public Concern?

Once a court has determined the speech has a sufficient connection to the school, it must assess if the speech is political, religious, or implicates some other matter of public concern. Currently, no precise test exists for determining if the challenged speech can be classified as such. In this inquiry, a court should assess whether the speech: (a) "is a subject of legitimate news interest; that is, a subject of general interest and of value and concern to the public;"[153] (b) "involves a matter of interest to the community;"[154] or (c) addresses "matters concerning government policies."[155] If the challenged speech implicates any one of these factors, it should be regarded as involving a matter of public concern.

Because of the greater burden of justification surrounding speech that implicates such matters, student speech that receives this classification may only fall within the school's regulatory authority if it is among the behaviors the Court has previously deemed outside the scope of First Amendment protection or if it falls into the narrow categories of speech the Court has expressly indicated are within the scope of school's power to regulate.[156] Accordingly, if the speech implicates a matter of public concern, a court should proceed to cross-reference the contested speech against the categories of speech the Supreme Court has recognized as unprotected speech in Table A, *infra*, and the few categories of speech the Court has recognized as unprotected for students in schools in Table B, *infra*. If the speech does not fall within one of the categories in Tables A or B, a court must dismiss the case in favor of the student because allowing schools to regulate such expression would be antithetical to the First Amendment's objective of affording citizens freedoms of inquiry and thought without governmental interference.[157]

---

151. B.L. *ex rel.* Levy v. Mahanoy Area Sch. Dist., 964 F.3d 170, 188-89 (3d Cir. 2020), *aff'd on other grounds*, 141 S. Ct. 2038 (2021).

152. *Kowalski*, 652 F.3d at 577.

153. Snyder v. Phelps, 562 U.S. 443, 453 (2011) (quoting City of San Diego v. Roe, 543 U.S. 77, 83-84 (2004)).

154. Nixon v. City & Cnty. of Denver, 784 F.3d 1364, 1367 (10th Cir. 2015).

155. *San Diego*, 543 U.S. at 80.

156. Mahanoy Area Sch. Dist. v. B.L. *ex rel.* Levy, 141 S. Ct. 2038, 2047 (2021).

157. *See* Griswold v. Connecticut, 381 U.S. 479, 482-83 (1965).

If a court determines the speech does not implicate a matter of public concern, it should move on to Step 3 of the evaluation process.

3. Step 3: Did the Speech Overlap with the Categories of Speech the Court has Already Addressed Concerning the Scope of the School's Regulatory Authority?

At this step, a court will go through the student speech regulatory guidelines in Tables A through C, *infra.* The guidelines consist of a composite list of all the behaviors explicitly or implicitly regarded by the Court as within or outside the regulatory authority of the school. This step is akin to Step 3 of the SSA's five-step sequential evaluation process, at which a final determination of disability can be made if the claimant's impairment appears among the listings.[158]

In the guidelines included in Tables A through C, *infra*, courts will be presented with several categories of speech to cross-reference the challenged speech against. These categories are: (1) Recognized Categories of Unprotected Speech;[159] (2) Types of Speech Recognized by the Supreme Court as Within the Scope of Schools' Disciplinary Authority;[160] and (3) Types of Speech Suggested by the Supreme Court as Within the Scope of Schools' Disciplinary Authority.[161] If a court determines the speech matches one of the categories of speech listed in either Table A: Recognized Categories of Unprotected Speech or Table B: Types of Speech Recognized by the Supreme Court as Within the Scope of Schools' Disciplinary Authority, the inquiry is over, and the student's speech is not protected. Accordingly, a court should dismiss the case finding that the student's punishment did not violate their First Amendment free speech rights.

If a court determines the speech matches one of the categories of speech listed in Table C: Types of Speech Suggested by the Supreme Court as Within the Scope of Schools' Disciplinary Authority, a rebuttable presumption is formed that the student's speech is not protected. However, because such speech has only been suggested as being within the bounds of schools' disciplinary authority, a court should still proceed to Step 4 and allow the student a chance to rebut the presumption that their speech is unprotected from punishment by the school. If a court determines the speech does not fall within any of the categories of speech included in Tables A through C, *infra*, it should also move onto Step 4 of the evaluation process.

158. Barnhart v. Thomas, 540 U.S. 20, 24-25 (2003).
159. *See infra* Table A.
160. *See infra* Table B.
161. *See infra* Table C.

4. Step 4: Did the Speech Pose a "Reasonably Foreseeable" Risk of, or has it Already Produced, a Substantial Disruption to the Pedagogical Interests of the School?

At this final step of the evaluation, a court should consider whether a jury would conclude that the speech had a "reasonably foreseeable"[162] risk of reaching the school,[163] whether the speech was specifically targeted at members of the school community,[164] whether the speaker encouraged other students' participation,[165] whether a disruption actually occurred,[166] and if it did, whether it produced a substantial disruption.[167] If the speech satisfies any one of these criteria, it would render the speech within the school's zone of regulatory authority—making the student's discipline permissible. If, however, a court determines the speech did not pose a reasonably foreseeable risk of substantial disruption to the school, a court must dismiss the case in favor of the student.

This step encapsulates the hallmark *Tinker* "substantial disruption" test, as well as the "reasonable foreseeability" test the Second Circuit articulated in *Wisniewski.*[168] As the reasonable foreseeability and substantial disruption tests have been subject to much of the same scrutiny for vagueness and inconsistent outcomes, this step includes pointed questions based on cases from the lower courts.[169] This step further serves as a final catch-all for speech that has a significant impact on the school that may have slipped through the cracks in Steps 1 through 3, or that has not been previously expressed by the Court as a category of speech that the school may regulate due to the unique "characteristics of the school environment."[170]

B. Getting Rid of Unnecessary Red Tape: Eliminating the Consideration of the Geographic Origin of Student Speech

As indicated above, the Supreme Court's decision in *Mahanoy* leaves many open questions concerning the parameters of on- versus off-campus speech for students, school administrators, and courts.[171] By placing emphasis on the location student speech originates from when evaluating schools' disciplinary authority, the Court in *Mahanoy* created a significant risk of

---

162. Wisniewski v. Bd. of Educ. of Weedsport Cent. Sch. Dist., 494 F.3d 34, 38-39 (2d Cir. 2007) (quoting Tinker v. Des Moines Indep. Cmty. Sch. Dist., 393 U.S. 503, 513 (1969)).

163. *Id.*

164. Mahanoy Area Sch. Dist. v. B.L. *ex rel.* Levy, 141 S. Ct. 2038, 2045 (2021).

165. *See* Doninger v. Niehoff, 527 F.3d 41, 50 (2d. Cir. 2008).

166. *See Tinker*, 393 U.S. at 514.

167. *See id.*

168. *Wisniewski*, 494 F.3d at 38-39 (quoting *Tinker*, 393 U.S. at 513).

169. Shannon M. Raley, Note, *Tweaking* Tinker*: Redefining an Outdated Standard for the Internet Era*, 59 CLEV. ST. L. REV. 773, 776, 796-97 (2011); Larissa M. Lozano, Note, *A 'Substantial and Material' Refinement of* Tinker, 46 N.M. L. REV. 171, 172, 179-83 (2016).

170. Morse v. Frederick, 551 U.S. 393, 408 (2007).

171. *See supra* Section II.D.3

future cases yielding inconsistent results for nearly identical forms of expression—as the spatial-temporal confines of modern schools are so hard to identify.[172] Justice Breyer even noted in the majority opinion that "given the advent of computer-based learning," distilling a meaningful standard for what constitutes off-campus speech would be extremely challenging due to the numerous exceptions and carveouts needed to accompany such a rule.[173] Yet, if the justices at the highest court in the country cannot distill such a distinction, how can lower courts reasonably be expected to do so in any reasonable or predictable manner?

In the absence of a clear-cut rule, the Court in *Mahanoy* noted circumstances that "may" call for a school's authority to address off-campus speech—referencing "severe bullying," "harassment targeting particular individuals," and "threats aimed at teachers or other students."[174] However, by using the permissive "may" as opposed to the imperative "shall" or "must," the Court provides for the possibility that even in these more extreme circumstances, schools *still* might not be authorized to regulate a student's speech merely based on its geographic origin.[175] Thus, under this standard, a school may rightfully punish a student for tweeting offensive images of a classmate every day from homeroom, but not the student who posts similarly inflammatory images from their house after school each day at 5:00 PM.[176] The differing outcomes for such similar behaviors beg the question of how this framework promotes the teaching of manners and civility, often viewed as an imperative of American public schools.[177]

The continued reliance on an on/off-campus distinction further creates a logistical challenge for educators and school officials to determine when discipline of students is permissible. Amidst what seems to be a never-ending pandemic and a mounting youth mental health crisis, public schools face an incredible number of challenges in educating and protecting the well-being of students.[178] Yet, instead of being able to respond quickly to what would ordinarily be routine disciplinary decisions, school officials instead are expected to sift through "multiple First Amendment standards and assay the bounds of the 'school environment'" to determine if a student can be suspended from after school activities for a week after a weekend of online activity mocking another student.[179] These challenges, accompanying the use of an on/off-campus speech distinction, highlight the need for the abandonment of this approach. Thus, instead of arbitrarily using a speech's geographic origin as a threshold requirement for whether student speech can

172. Mahanoy Area Sch. Dist. v. B.L. *ex rel.* Levy, 141 S. Ct. 2038, 2046 (2021).

173. *Id.* at 2045.

174. *Id.*

175. *See* ANTONIN SCALIA & BRYAN A. GARNER, READING LAW: THE INTERPRETATION OF LEGAL TEXTS 112 (2012); *Mahanoy*, 141 S. Ct. at 2045.

176. Reply Brief for Petitioner, *supra* note 2, at *2-3.

177. Bethel Sch. Dist. v. Fraser, 478 U.S. 675, 681 (1986).

178. U.S. DEP'T OF EDUC., SUPPORTING CHILD AND STUDENT SOCIAL, EMOTIONAL, BEHAVIORAL, AND MENTAL HEALTH NEEDS 1, 3, 7 (2021), https://www2.ed.gov/documents/students/supporting-child-student-social-emotional-behavioral-mental-health.pdf [https://perma.cc/9ACW-7SCF].

179. Reply Brief for Petitioner, *supra* note 2, at *2-3.

be regulated, the Court should instead adopt an approach focusing on the effects of the speech at issue with a proximate cause test like those historically used to govern students' out-of-school conduct.[180] The following section provides an example of the application of this proposed test and how it better incorporates the impact of the contested speech into its consideration of whether discipline is permissible.

C. Applying the Proposed Test

This Note began with a fictional anecdote in which Student B is suspended from school for ten days for her creation of a meme of one of her classmates, Student A, in which she superimposed Student A's picture on a cartoon ogre, with the caption "Weird Fat Fugly Ogre." Because the test proposed by this Note has never been applied by a court, this section seeks to illustrate how the test would function as applied to the facts provided in this fictional anecdote.

1. Step 1: Did the Speech Have a "Sufficient Nexus" to the School?

First, engaging with the threshold question of whether the ogre meme had a "sufficient nexus"[181] with the school's pedagogical interests, a court would likely determine the meme did have a sufficient nexus to the school. Here, as the Fourth Circuit determined in *Kowalski*, a lower court would likely find that the meme's inclusion of Student A's picture and its rapid circulation among the student body constituted a "targeted attack on a classmate . . . in a manner sufficiently connected to the school environment" to create a substantial disruption with the school's ability to discipline and protect the rights of its' students.[182] In addition, the meme was created loosely while Student B was heading home from school,[183] if going to Starbucks with her friends after school is to be viewed as a quick detour on her way home. While not corresponding to all the considerations included within Step 1, the meme's use of Student A's face and subsequent circulation to nearly the entire student body in less than 24 hours makes it highly probable that a jury would find a sufficient nexus to the school based on a totality of the circumstances.

180. *Mahanoy,* 141 S. Ct. at 2059-60 (Thomas, J., dissenting) (discussing how courts in the late 19th century used a "'direct and immediate tendency' to harm" standard for governing students' off-campus conduct (quoting Lander v. Seaver, 32 Vt. 114, 120 (1859))).

181. Kowalski v. Berkeley Cnty. Schs., 652 F.3d 565, 577 (4th Cir. 2011).

182. *Id.* at 567.

183. Mahanoy Area Sch. Dist. v. B.L. *ex rel.* Levy, 141 S. Ct. 2038, 2054 (2021) (Alito, J., concurring).

2. Step 2: Did the Speech Implicate a Matter of Public Concern?

Based upon the lower court's determination that the meme was sufficiently connected to the school, the court would then assess if the speech was political, religious, or implicated some other matter of public concern. Here, there appear to be no such interests addressed by the challenged speech. The meme contains neither an illustration of attitudes toward contemporary or historical events nor expresses a point of view or commentary on a social or political policy.[184] It simply appears to be born out of juvenile sniping and cliquishness—not a commentary on a matter of public concern.

3. Step 3: Did the Speech Overlap with the Categories of Speech the Court has Already Addressed Concerning the Scope of the School's Regulatory Authority?

At this step, the court would first go through the table titled "Recognized Categories of Unprotected Speech" in Table A, *infra*. Defamation appears to be the only category of unprotected speech the meme might fall under. However, while the meme identifies Student B by reasonable implication through its incorporation of her picture, the accompanying caption "Weird Fat Fugly Ogre" clearly indicates an opinion, not a fact, and thus fails to meet the standard for defamation.[185]

Turning to the table entitled "Types of Speech Recognized by the Supreme Court as Within the Scope of School Disciplinary Authority" in Table B, *infra*, the court would next consider the similarity of the meme to the types of speech provided in the table. Here, the meme did not use "lewd, indecent, or vulgar speech;"[186] "promote illegal drug use;"[187] or "bear the imprimatur of the school"[188]—thus, it would not seem to fall within the categories of speech explicitly declared within the scope of school's regulatory power by the Court in *Bethel*, *Morse*, and *Hazelwood* respectively.[189]

Next, the court would turn to the final table, "Types of Speech Suggested by the Supreme Court as Within the Scope of School Disciplinary Authority," in Table C, *infra*, to see if the meme matched any of the types of

184. *The Cartoon Analysis Checklist*, TEACHINGHISTORY.ORG, https://teachinghistory.org/sites/default/files/2018-08/Cartoon_Analysis_0.pdf [https://perma.cc/7G35-RDPD] (last visited Oct. 7, 2022).

185. HARVEY A. SILVERGATE ET AL., FOUND. FOR INDIVIDUAL RTS. IN EDUC., FIRE'S GUIDE TO FREE SPEECH ON CAMPUS 137-38 (Greg Lukianoff & William Creeley eds., 2d ed. 2012), https://dfkpq46c1l9o7.cloudfront.net/wp-content/uploads/2014/02/FIRE-Guide-to-Free-Speech-on-Campus-2nd-ed.pdf [https://perma.cc/RQ5W-L67R].

186. Bethel Sch. Dist. No. 403 v. Fraser, 478 U.S. 675, 676 (1986).

187. Morse v. Frederick, 551 U.S. 393, 408 (2007).

188. Hazelwood Sch. Dist. v. Kuhlmeier, 484 U.S. 260, 273 (1988).

189. 478 U.S. at 676; 551 U.S. at 408; 484 U.S. at 273.

speech provided therein. Here, as indicated above, the offending meme clearly targeted Student A. Therefore, the court would need to explore whether this targeting amounted to the sort of "severe bullying or harassment" Justice Breyer indicated as within the regulatory interests of the school in *Mahanoy*.[190] As no definition is provided for what constitutes "severe," the meme would seem to require something "beyond typical name-calling or teasing" and constitute more relentless or consistent attacks directed toward the victim.[191] Here, the meme alone, while offensive, would almost certainly not meet this high threshold. Thus, unless more information existed about previous attacks launched by Student B at Student A, the court would likely move on to Step 4 to make a final determination about whether Student B's suspension for creating the meme violated her free speech rights.

4. Step 4: Did the Speech Pose a "Reasonably Foreseeable" Risk of, or has it Already Produced, a Substantial Disruption to the Pedagogical Interests of the School?"

Assuming the court concluded the meme did not constitute severe bullying or harassment at Step 3, here, the court would engage in final considerations of whether the speech posed a "reasonably foreseeable" risk to the pedagogical interests of the school.[192] Specifically, the court should evaluate whether a reasonable jury would conclude that the speech would reach the school,[193] whether the speech was specifically targeted at members of the school community,[194] whether the speaker encouraged other students' participation,[195] whether the disruption occurred,[196] and if it did, whether it had a substantial impact.[197]

In this case, the offending meme seemingly meets all the criteria to constitute a "reasonably foreseeable" risk of disruption; the question is whether such disruption is "substantial."[198] As indicated above, the meme was specifically targeted at Student A—it superimposes Student A's Facebook profile picture onto the meme. Further, Student B clearly encouraged other students' participation in the attack on Student A by posting the meme to Twitter and sharing it with her classmates—both those with her in Starbucks and those still at school. The meme's viral dissemination would make it highly probable that a reasonable jury would conclude that the speech would reach the school.

---

190. Mahanoy Area Sch. Dist. v. B.L. *ex rel.* Levy, 141 S. Ct. 2038, 2045 (2021).

191. Rachel Simmons, *Extreme Bullying*, TEEN VOGUE (Sept. 21, 2010), https://www.teenvogue.com/story/extreme-bullying [https://perma.cc/A54B-DENC].

192. Wisniewski v. Bd. of Educ. of Weedsport Cent. Sch. Dist., 494 F.3d 34, 38-39 (2d Cir. 2007) (quoting Tinker v. Des Moines Indep. Cmty. Sch. Dist., 393 U.S. 503, 513 (1969)).

193. *See id.*

194. Mahanoy Area Sch. Dist. v. B.L. *ex rel.* Levy, 141 S. Ct. 2038, 2045 (2021).

195. *See* Doninger v. Niehoff, 527 F.3d 41, 50 (2d. Cir. 2008).

196. *See Tinker*, 393 U.S. at 514.

197. *See id.*

198. *Wisniewski*, 494 F.3d at 38-39 (quoting *Tinker*, 393 U.S. at 513).

However, as indicated above, the record does not seem to indicate that this incident was more than an isolated attack against Student A. Further, while the spread of the meme affected Student A deeply and led to her ultimately developing an eating disorder, more information would be needed to conclude whether her reaction was due to her "unreasonabl[e] fragil[ity]," as "otherwise protected speech [does] not become punishable" simply by offending the "hypersensitive."[199]

Considering recent news stories concerning the adverse impacts of social media on youth mental health[200] and all the facts provided,[201] it seems more likely than not that a court would conclude that Student B's meme, while created from an off-campus location, fell within the school's zone of regulatory authority based upon its significant impact on Student A. Thus, it seems highly probable that the court would find that Student B's suspension did not violate her First Amendment free speech rights.

D. Justifying the Proposed Test

As illustrated in the sample application above, the systematic nature of the proposed test provides for a streamlined approach to evaluate the merits of student speech cases. Through its clearly articulated, sequential inquiry and accompanying guidelines, this test would both help students to better understand the bounds of their speech rights and provide lower courts with more clarity on how to adjudicate cases. While there was some ambiguity at the final step as to the likely outcome of the case, this simply illustrates the high bar to which judges would be held to ensure no more speech than necessary is deemed beyond the scope of First Amendment protection for students. It is also important to note that many cases like *Mahanoy* would likely be dismissed following Step 1 due to the challenged speech's insufficient connection to the school. Thus, the number of cases for which such a time and resource-intensive analysis would be required is almost certainly slim.

Notwithstanding the appeal of such a systematic approach, some courts, like the Third Circuit, still contend that schools' disciplinary authority must not extend to off-campus speech, as to do so would constitute an intrusion into the lives of students and infringe upon parental autonomy.[202] However, the clear distinction insisted upon by the Third Circuit is quite illusory considering the explosion of computer-based learning brought on by COVID-19 lockdowns and the near-constant use of digital technology by school-aged children.[203] Thus, the Third Circuit's approach invites cutting off the ability of schools to discipline students almost entirely, denying schools the ability

199. Brewington v. State, 7 N.E.3d 946, 969 (Ind. 2014).

200. *See, e.g.*, Monica Anderson et al., *supra* note 3; Georgia Wells et al., *supra* note 3.

201. *See supra* Section I.

202. ROSS, *supra* note 10, at 225.

203. *See* Benjamin Herold, *The Decline of Hybrid Learning for This School Year in 4 Charts*, EDUC. WEEK (Sept. 27, 2021), https://www.edweek.org/technology/the-decline-of-hybrid-learning-for-this-school-year-in-four-charts/2021/09 [https://perma.cc/WN4V-TFP4]; *see also* Mahanoy Area Sch. Dist. v. B.L. *ex rel.* Levy, 141 S. Ct. 2038, 2045-46 (2021).

to serve their core function of instilling the nation's children with the skills and values necessary for them to develop into adults.[204] In contrast, if the Court were to adopt this proposed test, which fully takes into account the hyper-connected world of 2022, the disparate treatment experienced by students resulting from courts varied interpretations of the current standard would be significantly lessened.[205]

This country has more than 130,000 public elementary and secondary school principals, approximately 30,000 state court judges, and 1,700 federal court judges.[206] With so many potential players involved in a student's challenge to a disciplinary action, the likelihood of variability in interpretations of the scope of schools' authority to discipline students for their speech is exceptionally high.[207]

IV. CONCLUSION

For the reasons stated above, the Supreme Court's decision in *Mahanoy* provides school officials and lower courts with an insufficient standard for the regulation of off-campus student speech. With the continued reliance on virtual schooling to varying degrees as the pandemic continues, the need for clear guidance on this subject is more critical than ever. Thus, the Court should adopt the test proposed in this Note, as it provides a more comprehensive standard of review that requires lower courts to engage in a standardized, systematic inquiry to determine whether a student's First Amendment free speech rights have been violated. Such a standard is necessary to ensure students receive uniform enjoyment and protection of their First Amendment rights.

204. *See* Brown v. Bd. of Educ., 347 U.S. 483, 493 (1954).

205. *See supra* Section II.D.3.

206. *Table 105.50. Number of Educational Institutions, by Level and Control of Institution: Selected Years, 1980-81 Through 2017-18*, NAT'L CTR. FOR EDUC. STAT., https://nces.ed.gov/programs/digest/d19/tables/dt19_105.50.asp?current=yes [https://perma.cc/8LKP-NUP3] (last visited Nov. 20, 2021); INST. FOR THE ADVANCEMENT OF THE AM. LEGAL SYS., FAQS JUDGES IN THE UNITED STATES 3 (2014), https://iaals.du.edu/sites/default/files/documents/publications/judge_faq.pdf [https://perma.cc/A8VP-2757].

207. *See The Uniform College Athlete Name, Image, or Likeness Act (2021): A Summary*, UNIF. L. COMM'N, https://www.uniformlaws.org/HigherLogic/System/DownloadDocumentFile.ashx?DocumentFileKey=7fa9099b-eab8-3033-f4cf-69b4f09aae65&forceDialog=0 [https://perma.cc/SCT5-PCTW] (last visited Oct. 7, 2022).

V. APPENDIX

A. Table A: Recognized Categories of Unprotected Speech

Type of Speech	Standard	Next Steps?
1. Obscenity	In *Miller v. California*, 413 U.S. 15 (1973), the Supreme Court articulated the following three-part test to define obscenity: a) "The average person, applying contemporary community standards [to] find that the work, taken as a whole, appeals to the prurient interest;" b) The work to "describe in a patently offensive way, sexual conduct specifically defined by the applicable state law;" and c) "The work, taken as a whole" to "lac[k] serious literary, artistic, political, or scientific value." If each of these prongs is met, the expression is unprotected by the First Amendment.[208] **Sufficient Examples:** • Depictions or descriptions of "sexual acts," "masturbation, excretory functions, [or] lewd exhibition[s] of the genitals;"[209] • Erotic expression that would "conjure up psychic stimulation."[210] **Insufficient Examples:** • Depictions of nudity absent a showing of the genitals of the persons portrayed.[211]	-If the standard is met, the speech is not protected. The court should issue a summary judgment decision in favor of the school. -If the standard is not met, proceed to Row 2 for the "Defamation" inquiry.
2. Defamation	In a concurring opinion in *Rosenblatt v. Baer*, 383 U.S. 75, 92 (1996), Justice Stewart explained that defamation suits provide a means of redress and "the protection of [one's] own reputation from unjustified invasion and wrongful hurt."[212]	-If the standard is met, the speech is not protected. The court should issue a summary judgment decision in favor of the school.

208. Miller v. California, 413 U.S. 15, 24 (1973).
209. *Id.* at 25.
210. Cohen v. California, 403 U.S. 15, 20 (1971).
211. Jenkins v. Georgia, 418 U.S. 153, 161 (1974).
212. Rosenblatt v. Baer, 383 U.S. 75, 92 (1996) (Stewart, J., concurring).

	Sufficient Examples: • Statements that assert facts (not opinions), that identify their victims either by name or reasonable implication and that are capable of being proven false.[213] **Insufficient Examples:** • "Public officials, political candidates or [other] public figures may not recover" for defamatory statements made about them concerning their official conduct "unless the statement was both false and made with 'actual malice.'"[214] • Private figures seeking to recover for defamatory statements made against them concerning matters of public concern, "unless the statement was both false and made knowingly or at least negligently."[215] • Mere possession of obscene materials in one's own home.[216]	-If the standard is not met, proceed to Row 3 for the "Fraud" inquiry.
3. Fraud	In *Central Hudson Gas & Electric Corp. v. Public Service Comm'n of New York*, 447 U.S. 557 (1980), the Supreme Court established the standard that commercial speech which is fraudulent, or misleading will not receive free speech protections.[217] **Sufficient Examples:** • Speech that may lead to consumer deception.[218] **Insufficient Examples:** • This category of protected speech is not inclusive of all false statements. The Court has reasoned that "some false statements are inevitable if there is to be an open and vigorous	-If the standard is met, the speech is not protected. The court should issue a summary judgment decision in favor of the school. -If the standard is not met, proceed to Row 4 for the "Incitement" inquiry.

213. SILVERGATE ET AL., *supra* note 185, at 137-38.

214. MAGGS & SMITH, *supra* note 23, at 1076 (citing N.Y. Times Co. v. Sullivan, 376 U.S. 254, 287-88 (1964)).

215. *See id.* (citing Gertz v. Robert Welch, Inc., 418 U.S. 323, 352 (1974)).

216. SILVERGATE ET AL., *supra* note 185, at 44.

217. MAGGS & SMITH, *supra* note 23, at 1133 (citing Cent. Hudson Gas & Elec. Corp. v. Pub. Serv. Comm'n of N.Y., 447 U.S. 557, 566 (1980)).

218. KILLION, *supra* note 33 (citing Illinois *ex rel.* Madigan v. Telemarketing Assocs., Inc., 538 U.S. 600, 612 (2003)).

	expression of views in public and private conversation" and thus protected by the First Amendment.[219]	
4. Incitement	In *Brandenburg v. Ohio*, 395 U.S. 444 (1969), the Supreme Court reasoned that while the First Amendment protects speech that advocates breaking the rules or law, it does not protect speech that is aimed at "inciting or producing imminent lawless action and is likely to. . . produce such action."[220] **Sufficient Examples:** • Speech that instills fear that "serious evil will result" if the speech is not inhibited and that poses a reasonably imminent fear of danger.[221] **Insufficient Examples:** • Speech that creates "fear of serious injury cannot alone justify suppression of free speech and assembly."[222] • Speech that is merely morally reprehensible but presents no imminent threat of harm.[223]	-If the standard is met, the speech is not protected. The court should issue a summary judgment decision in favor of the school. -If the standard is not met, proceed to Row 5 for the "Fighting Words" inquiry.
5. Fighting Words	In *Chaplinsky v. State of New Hampshire*, 315 U.S. 568 (1942), the Supreme Court defined fighting words as words "which by their very utterance inflict injury or tend to incite an immediate breach of the peace" and deemed such speech as outside the scope of the First Amendment's free speech protections.[224] **Sufficient Examples:** • "Personally abusive epithets, which when addressed to the ordinary citizen are … inherently likely to provoke a violent reaction."[225]	-If the standard is met, the speech is not protected. The court should issue a summary judgment decision in favor of the school. -If the standard is not met, proceed to Row 6 for the "True Threats" inquiry.

219. *Id.* (citing United States v. Alvarez, 567 U.S. 709, 718-19 (2012)).
220. Brandenburg v. Ohio, 395 U.S. 444, 447-48 (1969).
221. Whitney v. California, 274 U.S. 357, 375-76 (1927) (Brandeis, J., concurring).
222. *Id.*
223. *Id.*
224. Chaplinsky v. New Hampshire, 315 U.S. 568, 572 (1942).
225. Cohen v. California, 403 U.S. 15, 20 (1971).

	Insufficient Examples: • Speech that is merely "upsetting or arouses contempt."[226]	
6. True Threats	In *Virginia v. Black*, 538 U.S. 343 (2002), the Supreme Court re-affirmed its recognition of true threats as unprotected speech.[227] It interpreted "true threats [to] encompass those statements where the speaker means to communicate a serious expression of an intent to commit an act of unlawful violence to a particular individual or group."[228] **Sufficient Example:** • "Forms of intimidation that are most likely to inspire fear of bodily harm."[229] **Insufficient Example:** • Political Hyperbole[230]	-If the standard is met, the speech is not protected. The court should issue a summary judgment decision in favor of the school. -If the standard is not met, proceed to Row 7 for the "Speech Integral to Criminal Conduct" inquiry.
7. Speech Integral to Criminal Conduct	In *Giboney v. Empire Storage & Ice Co*., 336 U.S. 490 (1949), the Court declared that the freedom of speech rarely extends its protections to speech "used as an integral part of conduct in violation of a valid criminal statute."[231] **Sufficient Examples:** • Speech that constitutes the solicitation of criminal activity;[232] • "Offers or requests to obtain illegal material;"[233] • Impersonation of government officials.[234] **Insufficient Examples:** • Overly broad prohibitions of speech, banning not only speech that promotes unlawful conduct but also "all truthful publications of	-If the standard is met, the speech is not protected. The court should issue a summary judgment decision in favor of the school. -If the standard is not met, proceed to Row 8 for the "Child Pornography" inquiry.

226. KILLION, *supra* note 33 (citing Snyder v. Phelps, 562 U.S. 443, 458 (2011)).
227. Virginia v. Black, 538 U.S. 343, 359 (2003).
228. *Id.*
229. *Id.* at 363.
230. *Id.* at 359.
231. Giboney v. Empire Storage & Ice Co., 336 U.S. 490, 498-99 (1949).
232. KILLION, *supra* note 33 (citing United States v. Williams, 553 U.S. 285, 297-98 (2008) & *Alvarez*, 567 U.S. at 721).
233. *Id.*
234. *Id.*

	facts" about a matter of public concern.[235]	
8. Child Pornography	In *New York v. Ferber*, 458 U.S. 747 (1982), the Court recognized child pornography as an additional category of unprotected speech that is subject to content-based regulation.[236] **Sufficient Examples:** • "Works that visually depict sexual conduct by children below" the age specified by statute.[237] **Insufficient Examples:** • Depictions of sexual conduct which are not obscene and "do not involve live performance or photographic or other visual reproduction of live performances."[238]	-If the standard is met, the speech is not protected. The court should issue a summary judgment decision in favor of the school. -If the standard is not met, proceed to the next table in Appendix B.

B. Table B: Types of Speech Recognized by the Supreme Court as Within the Scope of Schools' Disciplinary Authority

Type of Speech	Illustration	Next Steps?
1. "Sexually Explicit," "Indecent," "Lewd," or "Vulgar" Speech	In *Bethel Sch. Dist. No. 403 v. Fraser,* 478 U.S. 675, 685 (1986), the Supreme Court held that the First Amendment did not bar schools from disciplining students for using "offensively lewd" and indecent speech" in an assembly, stating "it is a highly appropriate function of public school education to prohibit the use of vulgar language."[239]	-If the standard is met, the speech is not protected. The court should issue a summary judgment decision in favor of the school. -If the standard is not met, proceed to Row 2.

235. *Giboney*, 336 U.S. at 498-99.

236. MAGGS & SMITH, *supra* note 23, at 1130 (citing New York v. Ferber, 458 U.S. 747, 765 n.18 (1982)).

237. *Ferber*, 458 U.S. at 764.

238. *Id.* at 765.

239. Bethel Sch. Dist. No. 403 v. Fraser, 478 U.S. 675, 676 (1986).

2. Speech Promoting Illegal Drug Use	In *Morse v. Frederick*, 551 U.S. 393 (2007), the Supreme Court held that the "'special characteristics of the school environment,' . . . and the governmental interest in stopping student drug abuse . . . allow schools to restrict student expression that they reasonably regard as promoting illegal drug use," and therefore deemed a principal's suspension of students for unfurling a banner that read "BONG HiTS 4 Jesus" during an approved out-of-school event as constitutional. [240]	-If the standard is met, the speech is not protected. The court should issue a summary judgment decision in favor of the school. -If the standard is not met, proceed to Row 3.
3. Speech Bearing the Imprimatur of the School	In *Hazelwood School District v. Kuhlmeier*, 480 U.S. 260 (1988), the Court held that the school officials' decision to withhold publication of student-written newspaper articles did not violate the student's First Amendment rights.[241] It further held that "other expressive activities . . . members of the public might reasonably perceive to bear the imprimatur of the school" were within the permissible scope of schools' disciplinary authority. [242]	-If the standard is met, the speech is not protected. The court should issue a summary judgment decision in favor of the school. -If the standard is not met, proceed to the next table in Appendix C.

C. Table C: Types of Speech Suggested by the Supreme Court as Within the Scope of Schools' Disciplinary Authority

Type of Speech	Next Steps?
1. All speech made during times when the school is responsible for the student[243]	-If the standard is met, the speech is not protected. The court should issue a summary judgment decision in favor of the school. -If the standard is not met, proceed to Row 2.
2. All speech taking place over school laptops, on the school's website, or through school email accounts or phones[244]	-If the standard is met, the speech is not protected. The court should issue a summary judgment decision in favor of the school. -If the standard is not met, proceed to Row 3.

240. Morse v. Frederick, 551 U.S. 393, 408 (2007).
241. Hazelwood Sch. Dist. v. Kuhlmeier, 484 U.S. 260, 273 (1988).
242. *Id.*
243. Mahanoy Area Sch. Dist. v. B.L. *ex rel.* Levy, 141 S. Ct. 2038, 2045 (2021).
244. *See id.*

3. All speech that takes place during extracurricular activities, including team sports and activities taken for school credit[245]	-If the standard is met, the speech is not protected. The court should issue a summary judgment decision in favor of the school. - If the standard is not met, proceed to Row 4.
4. All "speech taking place during remote learning"[246]	-If the standard is met, the speech is not protected. The court should issue a summary judgment decision in favor of the school. -If the standard is not met, proceed to Row 5.
5. Speech published by the school that is "poorly written, inadequately researched, biased, or prejudiced"[247]	-If the standard is met, the speech is not protected. The court should issue a summary judgment decision in favor of the school. -If the standard is not met, proceed to Row 6.
6. Speech that is deemed "unsuitable for mature audiences"[248]	-If the standard is met, the speech is not protected. The court should issue a summary judgment decision in favor of the school. -If the standard is not met, proceed to Row 7.
7. "Severe bullying or harassment" targeting others in the school community[249]	-If the standard is met, the speech is not protected. The court should issue a summary judgment decision in favor of the school. -If the standard is not met, proceed to Row 8.
8. "Threats aimed at teachers or other students" [250]	- If the standard is met, the speech is not protected. The court should issue a summary judgment decision in favor of the school. - If the standard is not met, proceed to Row 9.
9. Failure to adhere to school codes of conduct or "following rules concerning lessons" or the participation in school activities[251]	- If the standard is met, the speech is not protected. The court should issue a summary judgment decision in favor of the school. - If the standard is not met, proceed to Row 10.

245. *See id.*
246. *Id.*
247. Hazelwood Sch. Dist. v. Kuhlmeier, 484 U.S. 260, 271 (1988).
248. *Id.*
249. *Mahanoy*, 141 S. Ct. at 2045.
250. *Id.*
251. *Id.*

10. "Breaches of School Security Devices" [252]	- If the standard is met, the speech is not protected. The court should issue a summary judgment decision in favor of the school. -If the standard is not met, the speech is protected. The court should issue a decision in favor of the student.

252. *Id.*